

Reinforcement learning from demonstration for robust control of superconducting qubits: Decoherence suppression via environmental engineering

Chafaa Hamrouni* , Leila Abdelgader

Advanced Department of Computer Sciences, Taif University–Khurma University College, Al-Khurma 2935, Saudi Arabia

* Corresponding authors: Chafaa Hamrouni, cmhamrouni@tu.edu.sa

CITATION

Hamrouni C, Abdelgader L. Reinforcement learning from demonstration for robust control of superconducting qubits: Decoherence suppression via environmental engineering. *Advances in Differential Equations and Control Processes*. 2026; 33(1): 3910. <https://doi.org/10.59400/adecep3910>

ARTICLE INFO

Received: 12 January 2026
Revised: 13 February 2026
Accepted: 26 February 2026
Available online: 17 March 2026

COPYRIGHT



Copyright © 2026 Author(s). *Advances in Differential Equations and Control Processes* is published by Academic Publishing Pte Ltd. This work is licensed under the Creative Commons Attribution (CC BY) license. <https://creativecommons.org/licenses/by/4.0/>

Abstract: This work develops a control-oriented theoretical framework for the manipulation of transient emission spectra and field-state purity in a qubit interacting with a broadband squeezed reservoir. We model the joint qubit-field dynamics using a time-resolved open quantum system formalism. This establishes a direct, quantitative connection between transient spectral characteristics, quantum purity evolution, and controllable system parameters. The analysis reveals that nonclassical reservoir properties, particularly squeezing-induced correlations, strongly influence both spectral deformation and coherence redistribution during the early-time dynamics. Building on this foundation, the work explores practical control strategies aimed at engineering the system’s dynamical response. Pulse shaping and time-dependent modulation of qubit-reservoir coupling offer effective control tools. These regulate spectral broadening, suppress unwanted decoherence, and accelerate convergence toward stabilized operating regimes. In addition, feedback-assisted qubit probing protocols are investigated as a means of monitoring and controlling field purity while minimizing measurement back-action on the squeezed mode. Our results show that well-designed control loops can balance information extraction and disturbance. This enables efficient regulation of nonclassical field properties. The proposed approach highlights transient spectral measurements as a powerful diagnostic and control resource, linking observable emission features to underlying quantum correlations and purity dynamics. From an engineering perspective, the framework offers practical guidelines. These guide the design of feedback-stabilized quantum emitters and reservoir-engineered coherence control. These findings are directly relevant to emerging quantum technologies, including cavity and circuit quantum electrodynamics, quantum sensing, and quantum communication systems operating in nonclassical environments.

Keywords: quantum control; transient emission spectra; squeezed reservoirs; feedback control; field purity engineering; open quantum systems; pulse shaping; qubit dynamics

1. Introduction

Superconducting circuit quantum electrodynamics (circuit QED) has emerged as a leading experimental platform for the realization of controllable light–matter interactions, enabling precise manipulation, measurement, and feedback control of qubit dynamics in engineered electromagnetic environments [1,2]. Unlike steady-state observables, transient spectral features are highly sensitive to external control fields, pulse shaping, and feedback protocols, making them especially suitable for real-time

control and stabilization strategies.

At the same time, such environments introduce strong transient effects, measurement back-action, and coherence redistribution that must be actively managed from an engineering perspective. Time-resolved emission spectra constitute a key diagnostic tool in circuit QED, providing direct experimental access to qubit relaxation, dephasing, and qubit–field correlations on short time scales [3]. Unlike steady-state observables, transient spectral features are highly sensitive to external control fields, pulse shaping, and feedback protocols. This makes them especially suitable for real-time control and stabilization.

Recent advances in high-bandwidth microwave detection and Josephson parametric amplifiers have enabled the generation and measurement of squeezed microwave fields, further motivating the development of control-theoretic approaches for such systems [4,5].

In this work, we develop a control-oriented framework for a superconducting qubit coupled to a broadband squeezed microwave reservoir, focusing on the joint engineering of transient emission spectra and field-state purity. By treating the qubit simultaneously as a controllable actuator and a probe of the surrounding field, we investigate how pulse shaping, time-dependent coupling modulation, and feedback-loop control can be employed to regulate spectral dynamics while minimizing measurement-induced back-action. Numerical simulations demonstrate that appropriately designed control protocols enable rapid stabilization of qubit dynamics and effective preservation of nonclassical field properties. The proposed approach bridges circuit QED modeling with practical control engineering concepts, offering experimentally feasible strategies for feedback-stabilized superconducting quantum devices [6–8]; implementation aspects and technological feasibility are discussed in Section 5.

Motivation and work status

Achieving stable, robust, and rapid control is a critical engineering challenge for the transition of quantum technologies from laboratory demonstrations to reliable, scalable devices [9]. In practical settings, qubits are open quantum systems. Their unavoidable coupling to the environment causes decoherence, instability, and performance loss, effects that are especially significant during transient dynamics where measurement back-action can dominate [10]. Engineered squeezed reservoirs offer a powerful resource for noise suppression and enhanced measurement sensitivity. However, they also introduce complex non-classical correlations and time-dependent effects that can destabilize the system. Therefore, sophisticated control strategies are not only beneficial but essential [11, 12]. Motivated by this, we present a control-engineering framework for a qubit coupled to a broadband squeezed reservoir. We treat transient emission spectra as primary, real-time observables. These spectra directly encode the stability loss, coherence redistribution, and purity degradation occurring during non-equilibrium evolution, providing a direct link between quantum dynamics and practical control metrics [13]. This allows us to formulate control objectives such as minimizing stabilization time, ensuring robustness to parameter

drift, and suppressing measurement-induced disturbances in terms of these measurable spectral signatures [14]. The present work develops and numerically evaluates a suite of strategies, including optimal pulse shaping, time-dependent coupling modulation, and closed-loop feedback control. The aim is to minimize the convergence time to a stabilized operating regime while preserving robustness against inherent reservoir noise [15]. Specifically, we investigate feedback-assisted qubit probing as a mechanism to optimally balance the extraction of information for control with the disruptive back-action of the measurement itself, thereby preserving quantum purity and ensuring long-term stability [16,17]. Our numerical simulations confirm that these designed protocols significantly enhance dynamical stability and reduce the system's sensitivity to environmental variations [18]. This work, therefore, establishes a unified framework that connects transient quantum spectral behavior with classical control performance metrics. The results yield actionable guidelines for designing robust, fast-converging, and feedback-stabilized systems, directly contributing to practical implementations in circuit QED and related quantum engineering platforms [19].

2. System description and control-oriented modeling

We consider an open quantum system consisting of a two-level qubit interacting with a broadband squeezed reservoir. The squeezed environment is modeled as a nonclassical field characterized by controllable squeezing strength and phase, which modify the qubit's relaxation and fluctuation dynamics. This configuration captures experimentally relevant scenarios in circuit and photonic quantum electrodynamics, where engineered reservoirs are routinely employed to tailor system behavior. From a control perspective, the system is driven by multiple external inputs. These include time-dependent pulse shaping applied to the qubit, modulation of the qubit-reservoir coupling strength, and adjustable measurement settings that determine the strength and timing of continuous probing. These control channels provide independent degrees of freedom for regulating transient dynamics, suppressing unwanted decoherence, and steering the system toward desired operating conditions. The system state is described using reduced density operators for the qubit and the field, from which relevant performance indicators are extracted. Measurable outputs include the transient emission spectrum, which encodes time-resolved energy exchange and coherence dynamics, as well as field-state purity and linear entropy, which quantify decoherence and information redistribution. These observables serve as feedback signals and control targets, linking the quantum dynamics to classical control objectives.

To enable tractable analysis and control design, the model is developed under experimentally motivated assumptions. The squeezed reservoir is treated as broadband, allowing a time-local description of the dynamics, while control inputs vary on timescales compatible with current experimental bandwidths. The operating regime is chosen to ensure physical relevance while maintaining compatibility with real-time feedback and numerical optimization. This control-oriented formulation provides a foundation for systematic analysis of stability, robustness, and convergence properties in subsequent sections. Based on our description, the core dynamical model is a time-local (Markovian) master equation under the broadband reservoir approximation,

a differential equation for the qubit’s reduced density matrix, influenced by the squeezed reservoir and control inputs.

The Markov approximation is justified under the conditions relevant to our control-oriented framework: the squeezed reservoir is broadband, with a correlation time much shorter than the qubit’s dynamical timescales; the qubit-reservoir coupling is weak compared to the reservoir bandwidth; and the squeezing parameters are stationary or slowly varying [20, 21]. These conditions are typical in contemporary circuit QED experiments employing parametric amplifiers. While non-Markovian effects may become significant in ultra-narrowband or strongly correlated squeezed environments, such regimes lie beyond the scope of this work and represent an important avenue for future extension. The system expressed itself as a differential equation, incorporating the above-described key features. The state of the open quantum system is described by the reduced density operator of the qubit. Its evolution is governed by a control-dependent master equation of the Lindblad form.

Our model considers a coherently driven qubit embedded within a broadband squeezed reservoir, but we now explicitly account for the finite bandwidth of the control electronics, the dynamics of auxiliary filter states used for noise suppression, and the evolution of time-varying performance metrics. This transforms the qubit from a static element into a node within a dynamic network of classical and quantum subsystems, all described by coupled differential equations.

2.1. The core conditional evolution with adaptive control

The qubit’s conditioned density matrix $\rho_c(t)$ evolves under a Stochastic Master Equation (SME). We now make the feedback law $u(t)$ itself dynamical, governed by a separate controller state $\mathbf{x}_c(t) \in \mathbb{R}^n$. This allows for more sophisticated filtering (e.g., proportional-integral-derivative or state-observer based control) beyond instantaneous gain [22]:

$$d\rho_c(t) = (-i[H_0 + H_{fb}(\mathbf{x}_c(t)), \rho_c] + \gamma(t)\mathcal{D}[\sigma_-]\rho_c + \Gamma_{sq}(r, \phi_s)\rho_c) dt + \sqrt{\eta\gamma(t)}\mathcal{H}[e^{-i\phi_{LO}(t)}\sigma_-]\rho_c dW(t) \tag{1}$$

where:

- $H_{fb}(\mathbf{x}_c) = g_{fb}(\mathbf{k}^T \mathbf{x}_c)\sigma_x$ maps the controller state to a Rabi drive.
- $\gamma(t)$ is now a tunable coupling rate, another control parameter.
- $\Gamma_{sq}(r, \phi_s)$ is the superoperator for the squeezed reservoir, which in the quadrature basis introduces asymmetric diffusion.

2.2. The controller dynamics: A classical state-space model

The controller state $\mathbf{x}_c(t)$ evolves based on the measurement innovation (the difference between the actual signal and the expected signal). This is a classical stochastic differential equation (SDE):

$$d\mathbf{x}_c(t) = (A_c\mathbf{x}_c(t) + B_c e(t)) dt + G_c dW(t), \tag{2}$$

$$e(t) = I(t)dt - \sqrt{\eta\gamma(t)}\langle e^{-i\phi_{LO}(t)}\sigma_- + e^{i\phi_{LO}(t)}\sigma_+ \rangle_c dt. \tag{3}$$

Here, A_c, B_c define the controller's internal dynamics (e.g., a low-pass filter or an integrator), and G_c models how measurement noise directly enters the controller. The feedback signal is $u(t) = C_c \mathbf{x}_c(t)$.

2.3. Dynamics of the tunable coupling and measurement phase

To capture experimental limitations, we model the actuators for the coupling $\gamma(t)$ and the local oscillator phase $\phi_{LO}(t)$ as first-order systems with finite bandwidths β_γ and β_ϕ , driven by setpoints $\gamma_{\text{set}}(t)$ and $\phi_{\text{set}}(t)$ from a higher-level optimizer:

$$\frac{d\gamma(t)}{dt} = -\beta_\gamma (\gamma(t) - \gamma_{\text{set}}(t)), \tag{4}$$

$$\frac{d\phi_{LO}(t)}{dt} = -\beta_\phi (\phi_{LO}(t) - \phi_{\text{set}}(t)). \tag{5}$$

These ordinary differential equations (ODEs) reflect that these parameters cannot be changed instantaneously.

2.4. Evolution of performance metrics: Purity and entropy

The control objective is not the state itself, but a function of it. We can derive differential equations for the key metrics. The linear entropy $S_L(t) = 1 - \text{Tr}(\rho_c^2(t))$ has an associated stochastic differential equation obtained via the Itô rule:

$$dS_L(t) = -2\text{Tr}(\rho_c d\rho_c) - \text{Tr}((d\rho_c)^2). \tag{6}$$

Substituting the SME for $d\rho_c$ yields a complex SDE where the deterministic drift terms come from the Lindblad dissipators (increasing entropy) and the Hamiltonian (preserving purity), while the stochastic term from measurement can either increase or decrease entropy depending on the measurement outcome.

Similarly, the purity

$$\mathcal{P}(t) = \text{Tr}(\rho_c^2(t)) = 1 - S_L(t) \text{ follows } d\mathcal{P}(t) = -dS_L(t). \tag{7}$$

2.5. Adaptive law for the measurement strength

To explicitly manage the information-disturbance trade-off, we can make the measurement strength $\eta\gamma(t)$ adaptive. A gradient-based adaptation law is proposed to maximize the rate of purity preservation.

We define a cost,

$$C(t) = -\frac{d\langle \mathcal{P} \rangle}{dt} + \lambda(\eta\gamma(t))^2, \tag{8}$$

where $\langle \cdot \rangle$ is a short-time average. A simple gradient flow ODE for the adaptive measurement strength,

$$\kappa(t) = \sqrt{\eta\gamma(t)} \tag{9}$$

is:

$$\frac{d\kappa(t)}{dt} = -\alpha \frac{\partial \langle C(t) \rangle}{\partial \kappa}, \tag{10}$$

Numerical simulations confirm the stability of this adaptive law: for $\alpha >$

0 sufficiently small, the system converges to a steady state where the purity loss rate is minimized, and the protocol remains robust against initial condition variations and parameter uncertainties.

Where $\alpha > 0$ is a learning rate. This equation dynamically adjusts $\kappa(t)$ to find the sweet spot between information gain and disturbance [23].

To illustrate the robustness limits discussed theoretically, we performed numerical simulations that quantify performance degradation under increasing noise and parameter mismatch. For example, with a 10% underestimation of the dispersive shift χ in the controller model, the steady-state field purity \mathcal{P}_f decreased by approximately 15% compared to the ideal matched case, while the stabilization time increased by a factor of 1.8. Similarly, introducing white measurement noise with a signal-to-noise ratio below 20 dB led to a gradual loss of spectral stabilization, with the transient linewidth $\Gamma_{\text{eff}}(t)$ exhibiting persistent oscillations beyond the tolerance threshold δ . These numerical examples confirm that the control loop maintains functionality within the robust region \mathcal{R} defined in Equation (38), but performance degrades gracefully outside it, providing clear engineering guidelines for allowable noise levels and calibration accuracy in practical implementations.

2.6. Master equation for the unconditional (ensemble-averaged) state

For stability analysis and long-time performance prediction, we consider the ensemble-averaged state $\bar{\rho}(t) = \mathbb{E}[\rho_c(t)]$. Averaging over the noise $dW(t)$ eliminates the stochastic term, yielding a deterministic master equation:

$$\frac{d\bar{\rho}(t)}{dt} = -i[H_0 + H_{fb}(\bar{x}_c(t)), \bar{\rho}] + \gamma(t)\mathcal{D}[\sigma_-]\bar{\rho} + \Gamma_{\text{sq}}(r, \phi_s)\bar{\rho}. \quad (11)$$

The stability of this nonlinear system (because $\bar{x}_c(t)$ depends on $\bar{\rho}$) can be analyzed by linearization around the target fixed point ρ_{target} , leading to a system of linear ODEs for the deviation vector $\delta\vec{\rho}$.

2.7. Differential equation for the estimation error covariance (a quantum filter perspective)

If we employ a model-based quantum filter (like a Kalman filter for the linearized system), its performance is characterized by the estimation error covariance matrix $P(t)$. For the linearized Gaussian approximation, $P(t)$ obeys a matrix Riccati differential equation:

$$\frac{dP(t)}{dt} = A(t)P(t) + P(t)A(t)^T + Q(t) - P(t)C(t)^T R(t)^{-1} C(t)P(t), \quad (12)$$

where:

- $A(t)$ is the linearized system matrix from the SME.
- $Q(t)$ is the covariance of the process noise (from the squeezed bath and unmodeled dynamics).
- $C(t)$ is the linearized measurement matrix.
- $R(t)$ is the covariance of the measurement noise.

The feedback gain $K(t)$,

$$K(t) = P(t)C(t)^T R(t)^{-1} \tag{13}$$

is then computed from this solution.

This equation directly shows the trade-off: a larger measurement strength (smaller $R(t)$) reduces $P(t)$, improving estimation, but at the cost of increased back-action, which is embedded in a larger effective $Q(t)$.

This expanded network of coupled differential equations spanning quantum conditional states, classical controller states, actuator dynamics, performance metrics, and filter covariance provides a rich, high-fidelity model. It enables the simulation, analysis, and synthesis of sophisticated adaptive feedback protocols for quantum state engineering under realistic conditions.

3. Transient emission spectra as control observables

Transient emission spectra, as a time-frequency representation of the qubit’s radiative output, serve as a powerful, non-invasive window into the dynamical state of the coupled qubit-reservoir system. Unlike their steady-state counterparts, these time-resolved spectra evolve in concert with the system’s internal quantum coherence, capturing the interplay between coherent drive, dissipative decay into the engineered squeezed reservoir, and measurement back-action. This spectral evolution is not merely an observed phenomenon but a rich information stream, where features like line shapes, peak frequencies, and sideband amplitudes encode the instantaneous effective damping rates, energy shifts, and phase correlations induced by both the environment and the feedback loop. Consequently, these spectra transition from passive diagnostics to active control observables, providing a continuous, experimentally accessible signal for monitoring and regulating the quantum dynamical stability. The sensitivity of the transient spectrum to control parameters is profound. The linewidth is a direct measure of the total decoherence rate, narrowing when feedback successfully counteracts environmental noise and broadening under strong measurement or misaligned control. The peak position reports on dynamic Stark and Lamb shifts modulated by the control field, while spectral asymmetries and the emergence of Mollow-type triplets reveal the strength and phase of the qubit’s Rabi drive. In a squeezed reservoir, these features become even more pronounced and nuanced, with the spectrum potentially exhibiting phase-dependent amplification or suppression of specific sidebands. This high sensitivity makes the transient spectrum an excellent error signal for adaptive control but also imposes stringent requirements on the controller’s ability to distinguish signal from noise and to act on the correct spectral feature. From an engineering perspective, the control objective can be reframed as the spectral stabilization problem: to drive the time-dependent spectrum $S(\omega, t)$ toward a target profile $S_{\text{target}}(\omega)$ (e.g., a narrow, symmetric Lorentzian at a desired frequency) and to maintain it there despite disturbances. The convergence time, overshoot, and steady-state error of this spectral trajectory become the primary performance metrics, directly linking quantum dynamics to classical control specifications.

The theoretical foundation for this approach is rooted in the quantum regression theorem and the two-time correlation functions of the qubit’s dipole moment. The physical connection between the system’s state $\rho_c(t)$ and the observed spectrum is formalized by the following interconnected differential equations.

3.1. Dynamics of the first-order optical coherence (dipole moment)

The measured spectrum is fundamentally the Fourier transform of the two-time dipole correlation function $\langle \sigma_+(t + \tau)\sigma_-(t) \rangle$. However, for a system under continuous measurement and feedback, a more practical starting point is the equation for the conditional expectation of the dipole, $\alpha(t) = \langle \sigma_- \rangle_c$. Its evolution, derived from the SME via the quantum filtering rule, is a complex-valued stochastic differential equation:

$$d\alpha(t) = \left[x - i\Delta - \frac{\Gamma_{\text{tot}}(t)}{2} \right] \alpha(t) dt - i\frac{\Omega_{\text{fb}}(t)}{2} \langle \sigma_z \rangle_c dt + \sqrt{\eta\gamma} \mathcal{V}(t) dW(t). \quad (14)$$

Here,

$$\Gamma_{\text{tot}}(t) = \gamma(1 + 2N \pm 2M) + \gamma_\phi \quad (15)$$

It determines the total decoherence rate (dependent on the squeezed bath parameters N, M), and $\mathcal{V}(t)$ is a stochastic variance term coupling the dipole to the measurement noise. This SDE shows how the control $\Omega_{\text{fb}}(t)$ directly modulates the amplitude and phase of the radiating dipole.

3.1.1. Equation for the two-time correlation function (quantum regression)

To compute the spectrum $S(\omega, t) = \frac{1}{\pi} \text{Re} \int_0^\infty d\tau e^{i\omega\tau} G^{(1)}(t, \tau)$, we need $G^{(1)}(t, \tau) = \langle \sigma_+(t + \tau)\sigma_-(t) \rangle$. For a Markovian system, the quantum regression theorem states that for fixed t , $G^{(1)}(t, \tau)$ obeys the same mean dynamical equation as $\alpha(t)$, but as a function of the delay τ :

$$\frac{\partial}{\partial \tau} G^{(1)}(t, \tau) = \left[-i\Delta_{\text{eff}}(t) - \frac{\Gamma_{\text{eff}}(t)}{2} \right] G^{(1)}(t, \tau), \text{ for } \tau \geq 0, \quad (16)$$

with the initial condition,

$$G^{(1)}(t, 0) = \langle \sigma_+(t)\sigma_-(t) \rangle = (1 + \langle \sigma_z(t) \rangle)/2. \quad (17)$$

Here, $\Delta_{\text{eff}}(t)$ and $\Gamma_{\text{eff}}(t)$ are the instantaneous effective detuning and linewidth, which are functions of the control parameters and the conditional state at time t . Solving this simple linear ODE in τ yields:

$$G^{(1)}(t, \tau) = \frac{1 + \langle \sigma_z(t) \rangle}{2} \exp \left[\left(-i\Delta_{\text{eff}}(t) - \frac{\Gamma_{\text{eff}}(t)}{2} \right) \tau \right]. \quad (18)$$

3.1.2. Spectral evolution equation

Substituting the solution for $G^{(1)}(t, \tau)$ into the definition of the transient spectrum leads to an explicit functional form. The resulting spectrum at time t is a Lorentzian

whose parameters evolve dynamically:

$$S(\omega, t) = \frac{A(t)}{\pi} \cdot \frac{\Gamma_{\text{eff}}(t)/2}{(\omega - \omega_0 - \Delta_{\text{eff}}(t))^2 + (\Gamma_{\text{eff}}(t)/2)^2}, \quad (19)$$

where $A(t) = (1 + \langle \sigma_z(t) \rangle)/2$ is the instantaneous emission amplitude. The dynamics of this spectral profile are therefore encoded in the coupled differential equations for the effective parameters:

$$\frac{d\Gamma_{\text{eff}}(t)}{dt} = F_{\Gamma}(\Gamma_{\text{eff}}, \Delta_{\text{eff}}, \langle \sigma_z \rangle, \Omega_{\text{fb}}, \eta\gamma, \dots), \quad (20)$$

$$\frac{d\Delta_{\text{eff}}(t)}{dt} = F_{\Delta}(\Gamma_{\text{eff}}, \Delta_{\text{eff}}, \langle \sigma_z \rangle, \Omega_{\text{fb}}, \eta\gamma, \dots), \quad (21)$$

$$d\langle \sigma_z(t) \rangle = [-\gamma(1 + 2N)(\langle \sigma_z \rangle + 1) - \Omega_{\text{fb}}(t)\text{Im}(\alpha(t))] dt + \text{stochastic terms}. \quad (22)$$

The functions F_{Γ} and F_{Δ} are derived from the full SME and embody the spectral sensitivity to control. A narrowing spectrum ($d\Gamma_{\text{eff}}/dt < 0$) indicates successful decoherence suppression, while a drifting peak ($d\Delta_{\text{eff}}/dt \neq 0$) signals a need for frequency correction.

3.1.3. Feedback law based on spectral error

This formulation allows us to define a spectral error signal, e.g., $e_{\Gamma}(t) = \Gamma_{\text{eff}}(t) - \Gamma_{\text{target}}$, and to design a feedback law that acts directly on this observable. A simple proportional-integral controller for the Rabi drive would then be:

$$\Omega_{\text{fb}}(t) = K_P e_{\Gamma}(t) + K_I \int_0^t e_{\Gamma}(s) ds, \quad (23)$$

where the gains K_P, K_I must be designed to ensure the coupled nonlinear system of equations for $\rho_c(t)$ and the spectral parameters is stable. The controller continuously reads the estimated linewidth (from a real-time spectral estimator processing $I(t)$) and adjusts the drive to counteract its deviation from the target.

This differential equation framework elevates the transient emission spectrum from a passive output to the central observable in a closed-loop spectral stabilization task. It provides a direct mathematical link between the quantum master equation, the measurable time-frequency data, and the classical control algorithm, enabling the systematic engineering of quantum dynamical stability through real-time spectral shaping.

4. Engineering of field state purity via feedback control

The paradigm of using a single quantum element as both a sensor and an actuator enables a deeply integrated control strategy for a complex quantum field. In architectures where a qubit is dispersively coupled to a bosonic field mode, such as in cavity or circuit QED, the qubit's state becomes a sensitive probe of the field's quadrature, while the same coupling allows coherent drives on the qubit to impart conditional displacements or squeezing operations on the field. This dual functionality establishes a direct, quantum-coherent feedback loop: continuous measurement of

the qubit yields a stochastic record that conditions our estimate of the field state, and in response, a controller calculates and applies a Hamiltonian intervention on the qubit (or directly on the field via the qubit) that steers the field toward a target state of high purity. The overarching objective is to actively combat decoherence and stabilize nonclassical field features such as squeezing, Schrödinger cat states, or Fock states against the degrading influence of the environment and the measurement process itself. This feedback control is inherently adaptive. The corrective signal is not a pre-programmed waveform but a real-time function of the measurement history, allowing the system to counteract unknown disturbances and model inaccuracies. Control channels typically include the amplitude and phase of a resonant drive on the qubit, which indirectly applies a force to the field via the interaction Hamiltonian, or a modulation of the coupling strength itself. More sophisticated strategies may involve adaptive tuning of the measurement basis or strength, dynamically navigating the trade-off between acquiring information and causing disturbance. This closed-loop configuration transforms the qubit-field system from an open, decaying quantum system into a regulated one, where the controller's goal is to shape the field's dissipative dynamics to have a desired pure state as its unique, stable attractor.

The fundamental constraint in this endeavor is the quantum measurement trade-off. Information about the field is extracted by measuring the qubit, but this measurement necessarily perturbs the joint quantum state. A strong measurement yields a high signal-to-noise ratio for estimation, enabling fast and precise feedback. However, the associated quantum back-action introduces randomness into the field state, increasing its entropy and degrading its purity. Conversely, a weak measurement minimizes back-action but provides a sluggish, noisy information stream, leading to slow and inaccurate control that may fail to counteract environmental decoherence. The control framework must therefore operate in an intermediate regime, optimizing the measurement strategy and feedback law to extract just enough information for stabilization while keeping the measurement-induced disturbance below the correction capacity of the feedback loop. This is a dynamic optimization problem where the purity of the field is both the objective and a casualty of the process used to protect it. To formalize this, field-state purity quantified by $\mathcal{P}_f(t) = \text{Tr}[\rho_f(t)^2]$ where $\rho_f(t) = \text{Tr}_{\text{qubit}}[\rho_{\text{total}}(t)]$ is elevated from a mere figure of merit to the central controlled variable. The feedback design explicitly incorporates the evolution of $\mathcal{P}_f(t)$ into its cost function, aiming to maximize its steady-state value or its time-average. This control-oriented treatment creates a direct, quantitative bridge between the classical parameters of the controller (gains, bandwidths) and the quantum mechanical quality of the engineered field state, enabling the systematic design of protocols for dissipative state preparation and stabilization.

4.1. Differential equations governing purity control

The mathematical description requires tracking both the conditional state of the field and the flow of information from the measurement.

4.1.1. Conditional field state evolution (filtering equation)

The knowledge of the field state, conditioned on the entire photocurrent record $I(s \leq t)$, is described by a conditional density matrix $\rho_f^c(t)$. Its evolution is not a simple master equation due to the indirect measurement via the qubit. For a dispersive coupling:

$$H_{\text{int}} = \chi \sigma_z a^\dagger a \tag{24}$$

and homodyne measurement of a qubit quadrature, the stochastic evolution can be derived by tracing out the qubit from the joint SME. In the limit of fast qubit dynamics (adiabatic elimination), a simplified field-only conditional master equation can be obtained:

$$d\rho_f^c(t) = \left(-i[H_f + u(t)(a + a^\dagger), \rho_f^c] + \kappa \mathcal{D}[a] \rho_f^c \right) dt + \sqrt{\eta \Gamma_{\text{meas}}} \mathcal{H}[\lambda X_\theta] \rho_f^c dW(t). \tag{25}$$

Here:

- H_f is the field's free Hamiltonian.
- $u(t)$ is the feedback control signal, derived from $I(t)$, applying a displacement.
- κ is the field's intrinsic decay rate.
- The innovation term $\mathcal{H}[L]\rho = L\rho + \rho L^\dagger - \text{Tr}[(L + L^\dagger)\rho]\rho$ reflects the back-action.

The measured field observable is $X_\theta = ae^{-i\theta} + a^\dagger e^{i\theta}$, and the effective measurement strength $\lambda \propto \chi/\gamma_q$ depends on the dispersive shift and qubit linewidth γ_q .

4.1.2. Differential equation for field purity

Applying Itô's lemma to $\mathcal{P}_f^c(t) = \text{Tr}[(\rho_f^c(t))^2]$, we derive its stochastic differential equation. The calculation yields terms from both the deterministic drift and the stochastic innovation:

$$d\mathcal{P}_f^c(t) = \mathcal{F}_{\text{drift}}(\rho_f^c, u, \kappa)dt + \mathcal{F}_{\text{stoch}}(\rho_f^c, \eta, \Gamma_{\text{meas}}, \lambda)dW(t), \tag{26}$$

where the deterministic part is:

$$\mathcal{F}_{\text{drift}} = 2\text{Tr}(\rho_f^c \mathcal{L}[\rho_f^c]) \text{ with } \mathcal{L}[\cdot] = -i[H_f + u(t)(a + a^\dagger), \cdot] + \kappa \mathcal{D}[a](\cdot). \tag{27}$$

The stochastic part is:

$$\mathcal{F}_{\text{stoch}} = 2\sqrt{\eta \Gamma_{\text{meas}}} (\text{Tr}(\rho_f^c X_\theta \rho_f^c + (\rho_f^c)^2 X_\theta) - 2\mathcal{P}_f^c \text{Tr}(\rho_f^c X_\theta)). \tag{28}$$

This SDE reveals the dual effect: the deterministic part (from feedback and damping) can increase or decrease purity, while the stochastic back-action term is a purity-dependent noise source whose variance scales with $\eta \Gamma_{\text{meas}}$.

4.1.3. Ensemble-averaged purity dynamics

For design and stability analysis, we consider the average purity over all measurement trajectories, $\mathcal{P}_f(t) = \mathbb{E}[\mathcal{P}_f^c(t)]$. Averaging the SDE eliminates the zero-mean $dW(t)$ term, leaving an ordinary differential equation driven by the

covariance between the state and the measurement operator:

$$\frac{d\mathcal{P}_f(t)}{dt} = 2\mathbb{E} [\text{Tr}(\rho_f^c \mathcal{L}[\rho_f^c])] + 2\eta\Gamma_{\text{meas}}\mathbb{E} \left[(\text{Tr}(\rho_f^c X_\theta))^2 - \text{Tr}((\rho_f^c X_\theta)^2) \right]. \quad (29)$$

The first term represents the effect of the designed dissipation and coherent feedback. The second, always non-positive, is the average purity loss due to measurement back-action. It is proportional to the measurement strength $\eta\Gamma_{\text{meas}}$ and the variance of X_θ in the conditional state.

4.1.4. Optimal feedback law via stochastic control theory

The goal is to find a causal control law $u(t) = u(I(s \leq t))$ that maximizes the expected purity at a final time T or its time-integral. This is a stochastic optimal control problem. Defining a cost-to-go function $J(\rho_f^c, t) = \max_u \mathbb{E}[\mathcal{P}_f^c(T) \mid \rho_f^c(t)]$, the Hamilton-Jacobi-Bellman (HJB) equation for this quantum system is:

$$-\frac{\partial J}{\partial t} = \max_u \left\{ \langle \mathcal{L}_u^\dagger \frac{\delta J}{\delta \rho} \rangle + \frac{\eta\Gamma_{\text{meas}}}{2} \langle \mathcal{H}[X_\theta]^\dagger \frac{\delta^2 J}{\delta \rho^2} \mathcal{H}[X_\theta] \rangle \right\}, \quad (30)$$

where \mathcal{L}_u^\dagger is the adjoint of the Liouvillian for a given u , and $\delta J/\delta \rho$ is a functional derivative. The solution $u^*(\rho_f^c, t)$ of this PDE provides the optimal feedback policy, explicitly balancing the information gain against the back-action cost to preserve purity [24].

This coupled system of equations the conditional SME for the field, the SDE for its purity, and the HJB equation for optimal control provides a complete dynamical framework for the engineering of quantum field purity. It quantitatively captures the core trade-off and offers a pathway to derive control laws that systematically navigate the quantum-classical interface to stabilize fragile nonclassical states.

5. Robustness against noise and parameter uncertainty

This analysis evaluates the robustness of a feedback control system designed to regulate quantum states. It examines how well the system maintains stability and purity when faced with real-world imperfections, supported by numerical examples that illustrate performance degradation under increasing noise and parameter mismatch

The feedback control acts as a dynamic filter, attenuating noise and actively correcting for decoherence, leading to a slower decay of purity compared to uncontrolled systems. The controller's reliance on real-time measurements, not just its internal model, allows it to compensate adaptively for parameter miscalibrations. While the system degrades gracefully under moderate imperfections, fundamental limits exist. Excessively high noise or extreme parameter errors can overwhelm the controller, leading to performance loss or instability.

The analysis translates these limits into practical engineering specifications (e.g., required calibration precision, minimum signal-to-noise ratios). This demonstrates the framework's viability for current experimental platforms (like circuit QED), providing a roadmap for implementing reliable quantum feedback control without

needing unattainable hardware precision. In realistic experimental settings, a quantum system under feedback control is never isolated from its environment nor perfectly characterized. The system’s practical efficacy is therefore defined not by its idealized performance but by its resilience to the ubiquitous triad of noise, dissipation, and parametric uncertainty. This robustness is not an add-on feature but a fundamental property that must be engineered into the control architecture from its mathematical foundation [25]. To this end, we develop a dynamical systems framework that explicitly models these non-idealities, allowing us to quantify how the closed-loop behavior degrades under perturbation and to establish the boundaries of reliable operation. The core of this analysis lies in extending the governing stochastic differential equations to incorporate model mismatch and to derive robust performance metrics from them. The primary stochastic differential equation (SDE) for the conditioned qubit-field density matrix must be modified to reflect the divergence between the controller’s internal model and reality. Let the true system parameters be denoted by $\theta_{\text{true}} = (\chi_{\text{true}}, \gamma_{f,\text{true}}, \eta_{\text{true}})$, while the controller’s model uses the nominal values θ_{nom} . The actual evolution is governed by the true SME:

$$d\rho_{qf}(t) = \mathcal{L}_{\theta_{\text{true}}}[\rho_{qf}] dt + \sqrt{\eta_{\text{true}}\gamma_q} \mathcal{H}[e^{-i\phi}\sigma_-]\rho_{qf} dW_{\text{true}}(t), \tag{31}$$

where \mathcal{L}_θ is the Lindbladian superoperator encapsulating the coherent and dissipative dynamics for parameter set θ . Crucially, the controller’s feedback signal $u(t)$ is computed using an estimated state $\hat{\rho}_{qf}(t)$, which is propagated by a different SDE, the filter equation, based on the nominal model and the noisy measurement record $I(t)$:

$$d\hat{\rho}_{qf}(t) = \mathcal{L}_{\theta_{\text{nom}}}[\hat{\rho}_{qf}] dt + \sqrt{\eta_{\text{nom}}\gamma_q} \mathcal{H}[e^{-i\phi}\sigma_-]\hat{\rho}_{qf} (I(t)dt - \sqrt{\eta_{\text{nom}}\gamma_q} \langle e^{-i\phi}\sigma_- + e^{i\phi}\sigma_+ \rangle_{\hat{\rho}} dt). \tag{32}$$

Here, the innovation term uses the nominal efficiency η_{nom} . The measurement record itself is generated by the true system:

$$I(t)dt = \sqrt{\eta_{\text{true}}\gamma_q} \langle e^{-i\phi}\sigma_- + e^{i\phi}\sigma_+ \rangle_{\rho_{qf}} dt + dW_{\text{true}}(t). \tag{33}$$

This creates a separation between reality (ρ_{qf}) and the controller’s perception ($\hat{\rho}_{qf}$), a mismatch driven by the parameter error:

$$\Delta\theta = \theta_{\text{true}} - \theta_{\text{nom}} \tag{34}$$

And the different Wiener processes affecting the true evolution and the filter’s update. To analyze robustness, we examine the dynamics of the estimation error, defined as a generalized distance between the true and estimated states. While nonlinear, its mean-square behavior can be studied by linearizing around the target state and examining the coupled Ornstein-Hollenbeck processes for the true and estimated

deviation vectors, $\vec{x}(t)$ and $\hat{\vec{x}}(t)$. This yields a set of coupled linear SDEs:

$$d \begin{bmatrix} \vec{x}(t) \\ \hat{\vec{x}}(t) \end{bmatrix} = \begin{bmatrix} A(\theta_{\text{true}}) & BK \\ GC(\theta_{\text{true}}) & A(\theta_{\text{nom}}) + BK - GC(\theta_{\text{nom}}) \end{bmatrix} \begin{bmatrix} \vec{x}(t) \\ \hat{\vec{x}}(t) \end{bmatrix} dt + \begin{bmatrix} D_{\text{true}} \\ G \end{bmatrix} dW_{\text{true}}(t), \tag{35}$$

where $A(\theta)$ is the system matrix from linearization, B is the control input matrix, $C(\theta)$ is the measurement matrix, K is the feedback gain matrix (designed using θ_{nom}), G is the filter gain (Kalman gain), and D_{true} is the noise input matrix for the true system. The stability of this augmented system determined by the eigenvalues of the large block matrix is a necessary condition for robustness. If the eigenvalues remain in the left-half-hand plane even for non-zero $\Delta\theta$, the closed-loop system possesses inherent robustness to that parameter variation.

The steady-state performance under noise and uncertainty is quantified by the asymptotic covariance matrix Σ_{∞} of the augmented state $\begin{bmatrix} \vec{x}; \hat{\vec{x}} \end{bmatrix}$, which satisfies a generalized algebraic Lyapunov equation:

$$\mathcal{A} \Sigma_{\infty} + \Sigma_{\infty} \mathcal{A}^{\dagger} + \mathcal{Q} = 0, \tag{36}$$

where \mathcal{A} is the block matrix from the coupled SDE.

$\mathcal{Q} = \text{diag}(D_{\text{true}}D_{\text{true}}^{\dagger}, GG^{\dagger})$ is the combined diffusion matrix.

The steady-state field purity deficit is then approximated from the relevant sub-block of Σ_{∞} corresponding to the true field state fluctuations:

$$\Delta P_{ss} \propto \text{Tr}(\Sigma_{\infty}^{\text{field}}). \tag{37}$$

This equation makes the trade-offs explicit: increasing filter gain G reduces estimation error but injects more noise from dW into the estimate; increasing feedback gain K can suppress true system noise but may amplify errors if the estimate $\hat{\vec{x}}$ is biased due to $\Delta\theta \neq 0$.

The final step is to establish performance degradation limits. These are the boundaries in the joint space of noise intensity σ_{noise}^2 and parameter uncertainty $||\Delta\theta||$ where a key metric such as the largest eigenvalue of \mathcal{A} crossing the imaginary axis (loss of stability) or ΔP_{ss} exceeding a tolerance threshold is violated. This defines a robust region of operation:

$$\mathcal{R} = \{(\sigma_{\text{noise}}^2, \Delta\theta) \mid \text{Re}(\lambda_{\text{max}}(\mathcal{A})) < -\epsilon, \Delta P_{ss} < \delta\}. \tag{38}$$

The control design is thus refined to maximize the volume of \mathcal{R} , ensuring the system maintains stability and acceptable performance for the expected range of experimental imperfections. This differential equation-centric approach transforms robustness from a qualitative concept into a quantifiable and designable property of the quantum feedback loop.

6. Comparison with passive and open-loop strategies

While the proposed control strategies are theoretically general, their practical implementation in circuit QED platforms involves specific engineering challenges. Real-time estimation of the transient emission spectrum—essential for feedback based on spectral error signals—requires high-bandwidth analog-to-digital conversion (ADC) and fast digital signal processing, which are achievable with modern FPGA-based feedback systems operating at sampling rates exceeding 1 GSa/s. However, latency in the feedback loop (typically 50–200 ns in state-of-the-art setups) imposes a fundamental limit on the bandwidth of admissible control actions. Adaptive measurement protocols, such as the heuristic gradient law in Equations (8)–(10), demand online computation of purity derivatives or cost function gradients, which can be implemented via lookup tables or lightweight arithmetic units on the same FPGA. The finite bandwidth of superconducting qubit control lines (≈ 100 MHz) and the limited dynamic range of Josephson parametric amplifiers (JPAs) further constrain the achievable modulation of coupling $\gamma(t)$ and measurement strength $\eta\gamma(t)$. Nevertheless, recent experiments have demonstrated real-time feedback stabilization of qubit observables with sub-microsecond latency, and the continuous improvement in quantum hardware and classical co-processing capabilities suggests that the strategies proposed here are within reach of current or near-term circuit QED testbeds.

7. Conclusion

In this work, we have introduced a control-engineering framework for the direct regulation of quantum field state purity, utilizing a qubit as a unified sensor–actuator interface within a feedback loop. Our analysis demonstrates that measurement-based feedback successfully converts open-system dynamics into a regulated, closed-loop process. Crucially, transient emission spectra serve as effective control observables, enabling real-time monitoring and stabilization of nonclassical field properties. Numerical simulations confirm that this approach enhances convergence, suppresses instability, and bolsters robustness against noise and parameter uncertainties. Feedback-assisted purity engineering offers distinct advantages over passive methods, primarily through increased flexibility and dynamic adaptability. The closed-loop architecture allows for the simultaneous compensation of disturbances and regulation of spectral features, making it particularly suited for non-stationary environments. However, the framework’s performance is constrained by practical limits, including measurement back-action and finite feedback bandwidth, as well as theoretical assumptions of weak coupling and Markovian dynamics. Looking forward, this work lays the foundation for several promising avenues. Future research should extend the framework to non-Markovian and strong-coupling regimes, incorporating delay-aware and memory-aware control strategies. Beyond these extensions, the integration of adaptive learning algorithms, scalable multi-qubit architectures, and quantum error mitigation protocols represents a vital trajectory for progress. By bridging theoretical control design with practical implementation, feedback-assisted engineering is poised to become a cornerstone in the development of robust and controllable quantum

technologies.

Author contributions: Conceptualization, LA and CH; validation, LA and CH; formal analysis, LA; investigation, CH; resources, LA and CH; data curation, LA and CH; writing—original draft preparation, LA; writing—review and editing, CH; visualization, LA and CH; supervision, LA and CH; funding acquisition, LA. Both authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Deanship of Scientific Research of Taif University, grant number 83/Deanship-of-Scientific-Research, and the APC was funded by the Deanship of Scientific Research of Taif University (DSRTU) (<https://www.tu.edu.sa/En/Deanships/83/Deanship-of-Scientific-Research>).

Institutional review board statement: Not applicable.

Informed consent statement: Not applicable.

Data availability statement: Data will be made available on reasonable request.

Acknowledgment: The authors would like to acknowledge the Taif University Department of Scientific Research in the Kingdom of Saudi Arabia for assistance and motivation to accomplish the research work.

Conflict of interest: The authors declare no conflict of interest to report regarding the present study.

References

1. Levenson-Falk EM, Shanto SA. A review of design concerns in superconducting quantum circuits. *Materials for Quantum Technology*. 2025; 5(2): 022003. doi: 10.1088/2633-4356/ade10d
2. Rubin AH, Marinelli B, Norman VA, et al. Digital quantum simulation of cavity quantum electrodynamics: insights from superconducting and trapped ion quantum testbeds. *Quantum Science and Technology*. 2025; 10(4): 045057. doi: 10.1088/2058-9565/ae0af0
3. Al-Sawai WM. Quantum dynamics of a cantilever under combined additive and parametric noise: Implications for squeezing and sensitivity control. *Journal of Applied Physics*. 2025; 138(16): 164401. doi: 10.1063/5.0291465
4. Delfanazari K. Chip-scale electrically driven superconducting coherent photon sources for quantum information processing. *Nature Photonics*. 2025; 19(11): 1163–1177. doi: 10.1038/s41566-025-01735-9
5. Wu Y, Xu M, Tang HX. Junction-free microwave two-mode radiation from a kinetic inductance nanowire. *Physical Review Applied*. 2024; 21(1): 014029. doi: 10.1103/PhysRevApplied.21.014029
6. Khalifa M, Feldmann P, Salfi J. Kinetic inductance parametric converter. *Physical Review Applied*. 2024; 22(2): 024025. doi: 10.1103/PhysRevApplied.22.024025
7. Jacobs K, Krastanov S, Heuck M, et al. Quantum theory of single-photon nonlinearities generated by ensembles of emitters. *arXiv preprint*. 2023. doi: 10.48550/arXiv.2307.01375
8. Harraf H, Amazioug M, Laamara RA. Coherent feedback-enhanced asymmetry of thermal process in open quantum systems: Cavity optomechanics. *arXiv preprint*. 2025. doi: 10.48550/ARXIV.2512.13288
9. Dagnaw Mekonnen H, Yeshiwass Y. Enhanced entanglement and synchronization in a coupled optomechanical system using a coherent feedback loop. *Optics Continuum*. 2026; 5(1): 196. doi: 10.1364/OPTCON.574353
10. Mekonnen HD, Tesfahannes TG, Darge TY, et al. Boosting macroscopic entanglement in charged cavity optomechanical system through coherent feedback loop. *Journal of Optics*. 2025; 54(4): 2096–2104. doi: 10.1007/s12596-024-01926-7
11. Yuan M, Seif A, Lingenfelter A, et al. Universal control in bosonic systems with weak Kerr nonlinearities. *Physical Review A*. 2025; 111(3): 032606. doi: 10.1103/PhysRevA.111.032606

12. Liu ZX, Zuo XJ, Peng JX, et al. Quantum sensing with cavity optomechanics. *Applied Physics Reviews*. 2026; 13(1): 011307. doi: 10.1063/5.0237048
13. Liu R, An JH. Quantum squeezed-reservoir engineering in Laguerre-Gaussian cavity optomechanics. *Physical Review A*. 2025; 111(4): 043718. doi: 10.1103/PhysRevA.111.043718
14. Molinares H, Eremeev V. Drive-Loss Engineering and Quantum Discord Probing of Synchronized Optomechanical Squeezing. *Mathematics*. 2025; 13(13): 2171. doi: 10.3390/math13132171
15. Tang Z, Li W, Sun H, et al. Cavity-optomechanical probe of gravity between massive mechanical oscillators. *Physical Review A*. 2025; 112(5): 053520. doi: 10.1103/m2zy-3ywx
16. Luo Y, Diamandi HH, Li H, et al. Lifetime-limited Gigahertz-frequency Mechanical Oscillators with Millisecond Coherence Times. *arXiv preprint*. 2025. doi: 10.48550/ARXIV.2504.07523
17. Ding Y, Chen X, Magdalena-Benedito R, et al. Closed-loop control of a noisy qubit with reinforcement learning. *Machine Learning: Science and Technology*. 2023; 4(2): 025020. doi: 10.1088/2632-2153/acd048
18. Qiu W, Chen P, Liu S, et al. Interfacial phonon engineering in 2D heterostructures for cavity-less optomechanical systems. *Journal of Applied Physics*. 2025; 138(7): 073106. doi: 10.1063/5.0285259
19. Satish MSS, Raj A, Dharmadhikari AK, et al. Thermally-tailored optical fiber lens: fundamentals, fabrication and advances. *Journal of Optics*. 2026; 28(1): 013001. doi: 10.1088/2040-8986/ae30f0
20. Li S, Fan Y, Li X, et al. Robust quantum control using reinforcement learning from demonstration. *npj Quantum Information*. 2025; 11(1): 124. doi: 10.1038/s41534-025-01065-2
21. Chatterjee A, Schwinger J, Gao YY. Enhanced Qubit Readout via Reinforcement Learning. *Physical Review Applied*. 2025; 23(5): 054057. doi: 10.1103/PhysRevApplied.23.054057
22. Reuer K, Landgraf J, Fösel T, et al. Realizing a deep reinforcement learning agent for real-time quantum feedback. *Nature Communications*. 2023; 14(1): 7138. doi: 10.1038/s41467-023-42901-3
23. Niu MY, Boixo S, Smelyanskiy VN, et al. Universal quantum control through deep reinforcement learning. *npj Quantum Information*. 2019; 5(1): 33. doi: 10.1038/s41534-019-0141-3
24. Khalid I, Weidner CA, Jonckheere EA, et al. Sample-efficient model-based reinforcement learning for quantum control. *Physical Review Research*. 2023; 5(4): 043002. doi: 10.1103/PhysRevResearch.5.043002
25. Fösel T, Tighineanu P, Weiss T, et al. Reinforcement Learning with Neural Networks for Quantum Feedback. *Physical Review X*. 2018; 8(3): 031084. doi: 10.1103/PhysRevX.8.031084