

Spectral analysis of the Koopman operator recovers Hamiltonian parameters in open quantum systems

Jorge E. Pérez-García, Carlos Colchero, and Julio C. Gutiérrez-Vega*

*Photonics and Mathematical Optics Group,
Tecnológico de Monterrey, Monterrey, 64849, México.*

(Dated: December 5, 2025)

Accurate identification of Hamiltonian parameters is essential for modeling and controlling open quantum systems. In this work, we demonstrate that the multichannel Hankel alternative view of Koopman (*mHAVOK*) algorithm is a robust and reliable spectral data-driven method for retrieving Hamiltonian parameters from the evolution of first-moment observables in open quantum systems. The method relies on the discrete spectrum of the Koopman operator to obtain these parameters, which are computed using the *mHAVOK* algorithm; a theoretical connection to this affirmation is presented. The method is tested on noiseless quadratures of an open two-dimensional quantum harmonic oscillator and shown to retrieve oscillation frequencies, damping rates, nonlinear Kerr shifts, the qubit-photon coupling strength of a Jaynes-Cummings interaction, and the modulated frequency of a time-dependent Hamiltonian. The majority of the recovered parameters remained within 5% of their actual values. Compared with Fourier and matrix-pencil estimators, our approach yields lower errors for dynamics with strong dissipation. Overall, these findings suggest that Koopman operator theory provides a practical framework for studying quantum dynamical systems.

I. INTRODUCTION

In recent years, the understanding, characterization, and control of open quantum systems have become relevant topics in circuit and cavity quantum electrodynamics, among many other fields [1–3]. Having an accurate knowledge of oscillation frequencies, dissipation rates, and nonlinear interaction strengths is essential for the experimental calibration of quantum devices and for validating master-equation models. All these topics have also attracted considerable interest in the context of quantum computing [4, 5].

In general, parameters of interest are obtained from experimental data using frequency-domain spectroscopic algorithms or nonlinear curve fitting methods, which often require long acquisition times [6]. Techniques such as the Fourier, Prony, and matrix-pencil methods have also been used to extract frequencies and decay rates. However, their precision tends to decrease under strong damping conditions [7, 8]. More recently, machine learning and system identification methods have been proposed, but these algorithms often operate as black boxes, lacking physical interpretability and usually relying on large training datasets [9, 10]. Consequently, there is an increasing interest in fast, data-driven techniques that can directly infer Hamiltonian parameters from experimental data.

In this paper, we demonstrate that the novel multichannel Hankel alternative view of Koopman (*mHAVOK*) algorithm [11] is a robust and reliable spectral data-driven method for retrieving Hamiltonian parameters from the evolution of first-moment observables in open quantum systems. Such a technique relies on

analyzing the spectrum of the Koopman operator and improves the original HAVOK algorithm [12]. A detailed description of its functionality and performance can be found in Ref. [11].

With this considered, the main contributions of this work are summarized as follows:

(a) First, we demonstrate the theoretical connection between the Koopman operator spectrum and the *mHAVOK* model. To the best of our knowledge, this relationship has not been formally established in the literature. However, this step is essential, as it justifies why the discrete spectrum of the Koopman operator is retrieved by one of the dynamical matrices generated by the *mHAVOK* method.

(b) Second, we show that this spectrum accurately recovers the Hamiltonian parameters of open quantum systems. To test the method, we simulated the dynamics of an open two-dimensional Quantum Harmonic Oscillator (2D QHO) with multiple nonlinear effects and temporal perturbations to gather relevant data. Specifically, the algorithm was validated on the expected values of noiseless quadratures, retrieving oscillation frequencies, damping rates, nonlinear Kerr shifts, qubit-photon coupling strengths in a Jaynes-Cummings interaction, and the modulated frequency of a time-dependent Hamiltonian. The majority of the recovered parameters remained within 5% of their actual values. Compared with Fourier and matrix-pencil estimators, our approach yielded lower errors for dynamics with strong dissipation.

(c) Third, we contribute to reducing the demand for fast, data-driven techniques capable of inferring Hamiltonian parameters directly from experimental data. Since closed-form solutions for the master equations governing the evolution of quantum systems are often unavailable in real-world experimental settings, our method offers a data-driven, spectral-based approach to recover system parameters. For example, quadrature measurements

* juliocesar@tec.mx

could be obtained by homodyne detection [13] and subsequently analyzed using our framework, with the appropriate denoising filters applied.

The paper is organized as follows: Section II provides a mathematical justification on why the spectral analysis performed on *m*HAVOK is equivalent to studying the spectrum of the Koopman operator. In Sect. III, we briefly review some basic concepts of open quantum systems and the Lindblad master equation to establish notation and provide necessary formulas. The numerical results of the simulations and the operation of the *m*HAVOK method are included in Sect. IV for the case of both time-independent and time-dependent Hamiltonians. Section V summarizes the main contributions and limitations of this work. Finally, to keep this work relatively self-contained, we provide a brief overview of the *m*HAVOK method in Appendix A.

II. DEMONSTRATION OF THE KOOPMAN - MHAVOK CONNECTION

We begin by demonstrating the connection between the Koopman operator spectrum and the *m*HAVOK algorithm. This connection is crucial, as it provides the theoretical foundation for applying *m*HAVOK to the open quantum systems we are analyzing.

A. Concepts of Koopman operator theory

Consider a smooth dynamical system of the form $\dot{\mathbf{x}} = \mathbf{F}(\mathbf{x})$, with \mathbf{F} representing the vector field and \mathbf{x} the state vector. The Koopman semigroup is a linear, infinite-dimensional family of operators that advances in time functions of the state vector known as observables, not to be confused with the observables of quantum mechanics. Let $g \in \mathcal{H}$ represent an observable that lives in a Hilbert space \mathcal{H} and \mathcal{M} represent the state space, which we assume to be a smooth Riemannian manifold. Hence, the Koopman semigroup $(U^t)_{t \geq 0}$ acts on g as follows:

$$(U^t g)(\mathbf{x}) = g(\Phi^t(\mathbf{x})), \quad (1)$$

where Φ^t is the flow map. For a fixed sampling period $\Delta t > 0$, we define the discrete Koopman operator as $U^{\Delta t} := U$. Consequently, along a given trajectory $\mathbf{x}_k = \Phi^{k\Delta t}(\mathbf{x}_0)$, $\mathbf{x}_0 \in \mathcal{M}$ representing the initial condition, we have

$$g(\mathbf{x}_k) = (U^k g)(\mathbf{x}_0), \quad (2)$$

$$g(\mathbf{x}_{k+1}) = (Ug)(\mathbf{x}_k). \quad (3)$$

B. Connection of Koopman theory with delay embeddings

Suppose that we measure q observables $g_j \in \mathcal{H}$, with $j = \{0, 1, \dots, q-1\}$, and collect them into a vector-

valued observable $\mathbf{g} : \mathcal{M} \rightarrow \mathbb{C}^q$ of the form

$$\mathbf{g}(\mathbf{x}) := [g_0(\mathbf{x}), \dots, g_{q-1}(\mathbf{x})]^T. \quad (4)$$

We define a delay-embedded observable $\mathbf{h} : \mathcal{M} \rightarrow \mathbb{C}^{mq}$ as

$$\mathbf{h}(\mathbf{x}) := \begin{bmatrix} \mathbf{g}(\mathbf{x}) \\ \mathbf{g}(\Phi^{\Delta t}(\mathbf{x})) \\ \vdots \\ \mathbf{g}(\Phi^{(m-1)\Delta t}(\mathbf{x})) \end{bmatrix} = \begin{bmatrix} \mathbf{g}(\mathbf{x}) \\ (U\mathbf{g})(\mathbf{x}) \\ \vdots \\ (U^{m-1}\mathbf{g})(\mathbf{x}) \end{bmatrix}, \quad (5)$$

where m is known as the embedding dimension. Assuming N measurements are made along the trajectory, we define

$$\mathbf{g}_k := \mathbf{g}(\mathbf{x}_k) = \mathbf{g}(\Phi^{k\Delta t}(\mathbf{x}_0)), \quad (6)$$

as the k -th measurement, with k ranging from 0 to $N-1$. The corresponding delay-embedded measurements are then given by

$$\mathbf{h}_k := [\mathbf{g}_k, \mathbf{g}_{k+1}, \dots, \mathbf{g}_{k+(m-1)}]^T. \quad (7)$$

All \mathbf{h}_k are then stored in a block Hankel matrix,

$$\mathbf{H} = [\mathbf{h}_0 \quad \mathbf{h}_1 \quad \dots \quad \mathbf{h}_{N-m}]. \quad (8)$$

Given that the delay vectors in Eq. (7) are evaluations of the delay observable \mathbf{h} in Eq. (5) along the trajectory, we have

$$\mathbf{h}_k = \mathbf{h}(\mathbf{x}_k) = \mathbf{h}(\Phi^{k\Delta t}(\mathbf{x}_0)) = (U^k \mathbf{h})(\mathbf{x}_0). \quad (9)$$

Consequently, each column of \mathbf{H} corresponds to the consecutive action of the Koopman operator acting on \mathbf{h} and evaluated at the initial condition \mathbf{x}_0 ,

$$\mathbf{H} = [\mathbf{h}(\mathbf{x}_0) \quad (U\mathbf{h})(\mathbf{x}_0) \quad \dots \quad (U^{N-m}\mathbf{h})(\mathbf{x}_0)]. \quad (10)$$

The columns of \mathbf{H} in Eq. (10) form a Krylov sequence $\mathcal{F}_{N-m} = \{\mathbf{h}, U\mathbf{h}, \dots, U^{N-m}\mathbf{h}\} \subset \mathcal{H}$ whose columns span a corresponding Krylov subspace [14]. Similarly to the work in Ref. [14], we will assume the existence of \mathcal{K} , a finite-dimensional subspace of \mathcal{H} that is invariant under the action of the Koopman operator. Let d be the dimension of such subspace. Assuming the projection of \mathbf{h} onto \mathcal{K} is cyclic, i.e., repeated applications of the operator generate the whole invariant subspace, we have

$$\text{span}(\mathcal{F}_{N-m}) = \mathcal{K} \quad (11)$$

for $N-m \geq d-1$.

Next, a singular value decomposition (SVD) is performed on \mathbf{H} ,

$$\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \quad (12)$$

allowing for the existence of two orthonormal bases for the column and row space of \mathbf{H} , each given by the columns of \mathbf{U} and \mathbf{V} . If σ_i denotes the i -th singular value

and $v_i[k]$ the k -th component of the i -th right-singular vector, then each column of the block Hankel matrix can be expressed as

$$\mathbf{h}_k = \sum_{i=1}^{qm} \sigma_i v_i[k] \mathbf{u}_i := \sum_{i=1}^{qm} y_i[k] \mathbf{u}_i, \quad (13)$$

where

$$y_i[k] := \sigma_i v_i[k] \quad (14)$$

is the coordinate vector of \mathbf{h}_k in the basis of \mathbf{U} .

Subsequently, the first r singular vectors are retained in m HAVOK and assumed to be the leading SVD modes that best describe the system dynamics. A fixed value of r is now assumed; its optimal value is discussed in Subsection A.2. Let $\mathcal{H}_P = \text{span}\{\mathbf{u}_0, \dots, \mathbf{u}_{r-1}\} \subset \mathcal{H}$ denote the finite-dimensional subspace generated by the retained dynamical modes.

Reference [14] showed that for ergodic systems and in the long-time limit, the SVD of the Hankel matrix converges to the Proper Orthogonal Decomposition (POD) of the Krylov sequence. In our case, if $r \geq d$, then a subset of \mathcal{H}_P converges to an orthonormal basis of \mathcal{K} . Otherwise, if $r < d$ or if no exact finite-dimensional invariant subspaces exist, then \mathcal{H}_P is the best rank- r approximation to the span of the sampled Krylov sequence. In the following, we assume the existence of an $r \geq d$.

C. Projecting observables onto \mathcal{K}

Having discussed the assumptions and connections between the SVD modes and Krylov invariant subspaces, two orthogonal projection operators are defined on \mathcal{H}_P :

$$P : \mathcal{H}_P \rightarrow \mathcal{K}, \quad (15)$$

$$Q : \mathcal{H}_P \rightarrow \mathcal{H}_Q, \quad (16)$$

where $\mathcal{K} \oplus \mathcal{H}_Q = \mathcal{H}_P$ and $Q = I - P$, I representing the identity operator. Notice how P is a projection over the SVD-derived dominant modes that span $\mathcal{K} \subset \mathcal{H}_P$. Let $\mathbf{h}_u \in \mathcal{H}_P$ denote the projection of \mathbf{h} onto \mathcal{H}_P . Hence, \mathbf{h}_u is uniquely decomposed as

$$\mathbf{h}_u = P\mathbf{h}_u + Q\mathbf{h}_u := \mathbf{v} + \mathbf{u}, \quad (17)$$

with $\mathbf{v} \in \mathcal{K}$ and $\mathbf{u} \in \mathcal{H}_Q$. Considering the discrete-time Koopman operator acting on \mathcal{H}_P , its action can be decomposed as follows:

$$U = IUI = (P + Q)U(P + Q), \quad (18)$$

$$= PUP + PUQ + QUP + QUQ. \quad (19)$$

Thus, the projection operator P isolates the part of \mathbf{h}_u that lives in \mathcal{K} , while Q collects the residual component. The four blocks in Eq. (19) each describe the action of the operator that (i) stays within \mathcal{K} , (ii) maps residual dynamics onto the retained space ($\mathcal{H}_Q \rightarrow \mathcal{K}$), (iii) describes

how the retained dynamics interact with the orthogonal complement ($\mathcal{K} \rightarrow \mathcal{H}_Q$), and (iv) evolves entirely within the residual subspace.

If $QUP = 0$, then PUP gives an exact finite-dimensional restriction of the Koopman operator on that subspace. In the general case, however, QUP is different from zero, thereby being relevant to properly reconstruct the system dynamics, even if they are not captured in a Koopman-invariant subspace.

Interestingly, the above partition strongly resembles the Mori-Zwanzig formalism from statistical physics, which decomposes the system dynamics into resolved variables and unresolved components using projection operators, leading to closed equations of motion [15]. Indeed, a connection between Koopman theory and such formalism has already been reported in Ref. [16].

D. Regression-based approximation of the Koopman operator acting on \mathcal{K}

Having discussed the projection operators, the resolved part after one step is given by

$$P(U\mathbf{h}_u) = PUP\mathbf{v} + PUQ\mathbf{u}. \quad (20)$$

The m HAVOK algorithm constructs a forced linear model on the time series of the leading SVD coordinates, i.e., the first r right-singular vectors, in analogy to Eq. (20). The linear part of the model describes the dynamics that evolve under \mathcal{K} , while the forcing is assumed to come from its orthogonal projection.

Recalling the assumption that a subspace of \mathcal{H}_P spans \mathcal{K} , the method performs a regression step [Eq. (A5)] to classify the i -th main dynamical mode (out of the r ones) as linear or nonlinear. The linear modes are considered to span \mathcal{K} , while the nonlinear ones describe the unresolved dynamics.

Following Eq. (14), if r_c and r_f each denote the index sets of linear and nonlinear components, then the following coordinate vectors are defined:

$$\mathbf{y}_k := [y_j[k] : j \in r_c]^T \in \mathbb{C}^{|r_c|}, \quad (21)$$

$$\mathbf{u}_k := [y_j[k] : j \in r_f]^T \in \mathbb{C}^{|r_f|}, \quad (22)$$

with $|\cdot|$ representing the cardinality of the sets. With this considered, the algorithm models the dynamics as

$$\mathbf{y}_{k+1} = \mathbf{A}\mathbf{y}_k + \mathbf{B}\mathbf{u}_k, \quad (23)$$

where \mathbf{A} and \mathbf{B} are obtained from a least-squares regression. Considering \mathbf{y}_k and \mathbf{u}_k are coordinate vectors of $P\mathbf{h}_u$ and $Q\mathbf{h}_u$ in the $\{\mathbf{u}_j\}$ basis, \mathbf{A} and \mathbf{B} can be thought of as finite-dimensional approximations to the Koopman blocks

$$\mathbf{A} \approx PUP|_{\mathcal{K}}, \quad \mathbf{B} \approx PUQ|_{\mathcal{H}_Q}. \quad (24)$$

In this view, \mathbf{A} approximates the action of U on the invariant subspace spanned by the linear modes, while

\mathbf{B} quantifies how the nonlinear components influence the resolved coordinates.

Moreover, the spectrum of \mathbf{A} approximates the spectrum of the Koopman operator restricted to \mathcal{K} . Let

$$\sigma(\mathbf{A}) = \{\lambda \in \mathbb{C} \mid \det(\mathbf{A} - \lambda I) = 0\} \quad (25)$$

represent the set of retrieved eigenvalues λ_i . Even though the Koopman operator has both a continuous and discrete spectrum [17], \mathbf{A} recovers the latter one, as finding a finite-dimensional realization of the Koopman operator on a Koopman-invariant subspace is, in general, equivalent to the existence of the pure point spectrum on that subspace [14].

It is important to note, however, that the existence of a purely discrete spectrum on a dynamical system is largely a theoretical condition [18]. In practice, the majority of the systems exhibit both a continuous and discrete spectrum. *mHAVOK* overcomes such limitation by allowing the existence of \mathbf{B} , a (possibly) rectangular matrix that contains information of nonlinear components. Hence, we hypothesize that such matrix might have a relation with the continuous spectrum. Considering that the spectral analysis of \mathbf{A} alone is sufficient for a proper recovery of Hamiltonian parameters, as will be shown later, such a hypothesis is therefore not pursued in the current work.

III. PRELIMINARY CONCEPTS AND CONSIDERATIONS OF THE PROBLEM

In this section, we briefly recall some basic concepts of open quantum systems to establish notation and provide necessary formulas.

The dynamics of open quantum systems can be characterized by the Lindblad master equation, mainly in the Born and Markov approximations [19]. Its standard form reads as

$$\frac{d\rho}{dt} = -\frac{i}{\hbar}[H, \rho] + \sum_j \left(L_j \rho L_j^\dagger - \frac{1}{2} \{ L_j L_j^\dagger, \rho \} \right), \quad (26)$$

where ρ is the density matrix, H is the Hamiltonian, $\{\cdot, \cdot\}$ denotes the anti-commutator, L_j are jump operators that model system-environment interactions and \cdot^\dagger represents the complex conjugate. A comprehensive derivation of Eq. (26) can be found in Ref. [20]; for a deeper explanation, we refer the reader to Ref. [21].

Throughout this work, we model an open 2D QHO solving Eq. (26) under a common environment for both the x and y Cartesian components of the oscillator. The nature of the bath is described as follows:

The emission of photons from the system to the bath, also known as a decay, is given by

$$\mathcal{L}_c = a_x + e^{i\phi} a_y, \quad (27)$$

with ϕ being the relative phase between the two system modes and a_i representing the annihilation operator for the i -th component. Physically, the bath does

not recognize whether the emission came from the x or y component; it only couples both emissions to the linear combination in Eq. (27). On the other hand, absorption of photons from the bath to the system, a process also known as excitation, is given by \mathcal{L}_c^\dagger .

Additionally, the bath has a temperature $T > 0$ and is considered to be a collection of harmonic oscillators, from which the system can both emit and absorb quanta. The average number of photons in the bath at frequency ω_i is given by the Bose-Einstein distribution [21]

$$\bar{n}_i(\omega_i, t) = \left[\exp\left(\frac{\hbar\omega_i}{k_B T}\right) - 1 \right]^{-1}, \quad (28)$$

where $i = \{x, y\}$ and k_B is the Boltzmann constant. Taking this into account, the jump operators defined in Eq. (26) are given by

$$L_i^{\text{decay}} := \mathcal{L}_c \sqrt{\kappa(\bar{n}_i + 1)}, \quad (29)$$

$$L_i^{\text{excitation}} := \mathcal{L}_c^\dagger \sqrt{\kappa\bar{n}_i}, \quad (30)$$

with κ denoting the damping rate. Note that decay occurs even in the absence of photons, such as in a vacuum. These phenomena can be attributed to quantum fluctuations within the system [21].

IV. RETRIEVING HAMILTONIAN PARAMETERS

The spectral analysis performed to retrieve the Hamiltonian parameters using *mHAVOK* is illustrated in Fig. 1. We begin by generating the expected values for n observables of the open quantum system of interest using an accurate numerical solver of the Lindblad master equation [22]. This data is then used as input for the *mHAVOK* algorithm, which allows us to (1) retrieve a forced linear model of the system dynamics and (2) utilize the spectrum of the matrix \mathbf{A} to estimate the desired Hamiltonian parameters for the open quantum system.

In this work, we will consider a 2D QHO as our test system, incorporating various time-independent and time-dependent nonlinear effects. The initial state chosen for our simulations is a coherent state with an eigenvalue of $\alpha = 1$. The parameters used in these simulations are detailed in Table I. The evolution of the quadrature phase-space vector

$$\mathbf{x} = \begin{bmatrix} \langle x(t) \rangle \\ \langle y(t) \rangle \\ \langle p_x(t) \rangle \\ \langle p_y(t) \rangle \end{bmatrix} \quad (31)$$

is obtained by solving the Lindblad master equation Eq. (26) with the QuTiP solver [22]. This data is then provided as input to the *mHAVOK* algorithm. The user-defined threshold used in the algorithm for discerning linear from nonlinear components was set to $\tau = 0.95$. Furthermore, the embedding dimension was set to $m = 100$,

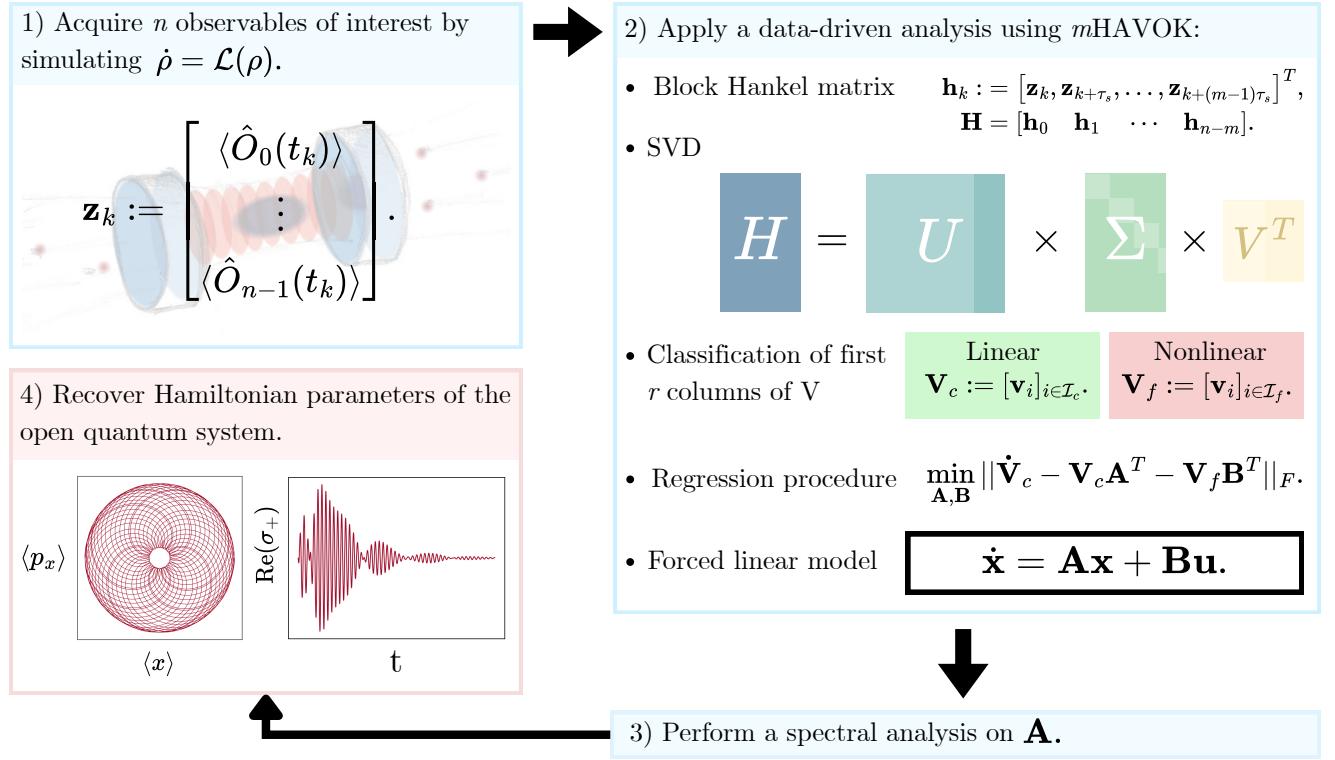


Figure 1. Overview of the pipeline employed in this work. (1) The expected values of observables are first obtained by simulating the Lindblad master equation. (2) Such data is then analyzed through the *mHAVOK* framework, where two dynamical matrices are retrieved. A spectral analysis is performed on one of them (3), allowing recovery of the Koopman generator’s discrete spectrum and, thus, the Hamiltonian parameters of the open quantum system (4).

Table I. Simulation parameters employed in this work.

Variable	Value
Simulation time	$t_f = 50$
Time step	$\Delta t = 0.01$
Oscillation frequencies	$\omega_x = 2\pi$ $\omega_y = \pi$
Environment temperature	$T = 2.0$
Damping rate	$\kappa = 0.1$
Fock-state truncations	$n_x = 10$ $n_y = 10$

consistent with previous research on the HAVOK method [11]. These parameters are part of the *mHAVOK* input configuration, all of which are further elaborated in Appendix A.

In the following, a hat over a variable indicates that the corresponding value was numerically recovered from the spectrum of \mathbf{A} . Conversely, if the variable lacks a hat, it represents the original parameter used to solve the Lindbladian. Hence, the percent error between the two quantities will be denoted as $e_{\lambda\hat{\lambda}}$. This metric is calculated and then compared with traditional methods for recovering oscillation frequencies, such as the Fast Fourier Transform (FFT) and the generalized pencil-of-

function (matrix pencil) method [7]. The matrix pencil method is an algorithm that effectively extracts damping factors and oscillation frequencies from an input signal, under the assumption that a sum of complex exponentials provides a suitable basis for it [8].

A. Two-dimensional quantum harmonic oscillator

The Hamiltonian of a 2D QHO is given by

$$H_0 = \hbar\omega_x a_x a_x^\dagger + \hbar\omega_y a_y a_y^\dagger, \quad (32)$$

where ω_i are the oscillation frequencies and a_i^\dagger denote the creation operators. For notational convenience, the zero of the energy has been shifted upwards by $\hbar(\omega_x + \omega_y)/2$, merely by redefining the reference of the energy. Throughout this work, the ladder operators and, hence, the quadratures are expressed in dimensionless units.

Considering that the Hamiltonian in Eq. (32) is quadratic and the jump operators in Eqs. (29) and (30) are linear, then the Heisenberg equations of motion for the first moments evolve linearly and have closed form expressions [21]. Consequently, the evolution of the quadrature vector Eq. (31) is given by the following ordinary differential equation (ODE):

$$\dot{\mathbf{x}} = \mathbf{A} \mathbf{x}, \quad (33)$$

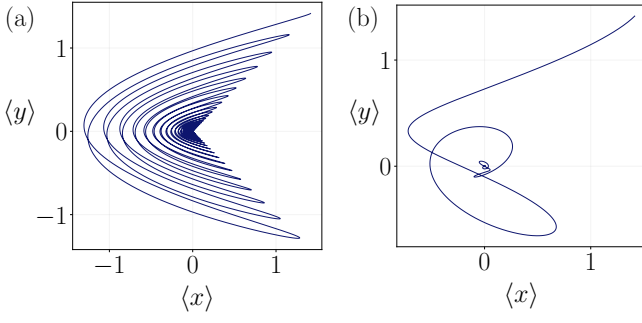


Figure 2. *mHAVOK* phase-space reconstruction of an open 2D QHO for the x and y components using a damping rate of (a) $\kappa = 0.1$ and (b) $\kappa = 1.0$. No nonlinearities were introduced.

where

$$\mathbf{A} = \begin{bmatrix} -\kappa/2 & \omega_x & 0 & 0 \\ -\omega_x & -\kappa/2 & 0 & 0 \\ 0 & 0 & -\kappa/2 & \omega_y \\ 0 & 0 & -\omega_y & -\kappa/2 \end{bmatrix}.$$

Note that \mathbf{A} depends on the oscillation frequencies and damping coefficients, but not on the bath occupations \bar{n}_i . In Fig. 2, we illustrate the system evolution for two different values of the damping rate κ and data taken from Table I. The subplot 2(a) shows the trajectory of the expected value of the particle in the (x, y) plane, where we see that the x -mode oscillates twice as fast as the y -mode, in agreement with the simulated oscillation frequencies. Furthermore, the system follows a parabolic Lissajous curve, with the amplitudes decaying and collapsing towards the origin due to the open setup. In contrast, Fig. 2(b) shows a substantially faster collapse towards the origin resulting from an order-of-magnitude increase in the damping rate. Consequently, no parabolic trajectories are identified.

Regarding the *mHAVOK* model, the optimal cutoff rank r_{opt} was found to be $r = 4$, which coincides with the number of input observables. This correspondence is not a general feature but rather a consequence of the system dynamics: since the input quadrature vector obeys the linear ODE in Eq. (33), higher-order modes do not contribute significantly to the reconstructed dynamics.

After determining the optimal cutoff rank, we obtained the retrieved dynamical matrix \mathbf{A} . By applying Schur decomposition [23], we found that this matrix takes the form of A as in Eq. (33). Furthermore, its four eigenvalues, up to numerical error, appeared in conjugate pairs of the form $-\kappa/2 \pm \hat{\omega}_i$. As discussed in Sect. II, these eigenvalues correspond to those of the Koopman operator restricted to the invariant subspace spanned by our observables, thereby providing a linear description of the system dynamics.

Once the spectrum of \mathbf{A} was obtained, the percent error between the retrieved Hamiltonian parameters and their true values was calculated. For the case in Fig. 2(a)

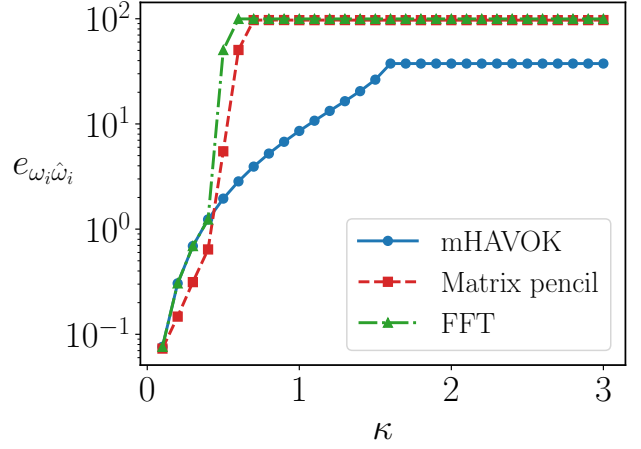


Figure 3. Percent error comparison among three different methods for determining the oscillation frequencies $\hat{\omega}_i$ of the 2D QHO. *mHAVOK* exhibits a slower increase in percent error compared to existing alternatives.

where $\kappa = 0.1$, the average percent error between ω_i and $\hat{\omega}_i$ was $e_{\omega_i \hat{\omega}_i} = 0.08\%$, whereas $e_{\kappa \hat{\kappa}} = 7.2 \times 10^{-4}\%$. On the other hand, when increasing 10 times the value of κ in the simulation setup [Fig. 2(b)], a notable increase is observed in the oscillation-frequency error, with $e_{\omega_i \hat{\omega}_i} = 8.58\%$. Interestingly, the error for the damping rate decreased almost four times once we increased its value, with $e_{\kappa \hat{\kappa}} = 1.8 \times 10^{-4}\%$. Figure 3 compares the percent error associated with three different methods for retrieving the oscillation frequencies as the damping rate increases. Although the matrix-pencil scheme is more effective for weak damping rates ($\kappa < 0.5$), *mHAVOK* significantly enhances the recovery for stronger rates.

B. Kerr nonlinearities

In classical optics, the Kerr effect appears in a nonlinear medium whose refractive index depends on the intensity of the electric field [24]. In quantum optics, this effect shifts the oscillation frequencies by an amount proportional to the number of photons present at a given time. The Hamiltonian that describes this phenomenon is given by [25]:

$$H_{\text{Kerr}} = \hbar \left(\frac{\chi_x}{2} a_x^{\dagger 2} a_x^2 + \frac{\chi_y}{2} a_y^{\dagger 2} a_y^2 + \chi_{xy} n_x n_y \right), \quad (34)$$

with $n_i = a_i a_i^{\dagger}$ denoting the number operator and χ_i representing Kerr strength coefficients. The resulting energy spectrum forms an anharmonic ladder [26], with the transition frequency between successive Fock states following $\omega_i + \chi_i n_i$, with n_i being a natural number.

The Hamiltonian of the tested system is then

$$H = H_0 + H_{\text{Kerr}}, \quad (35)$$

where H_0 is given by Eq. (32). Considering that Eq. (34) contains quartic terms like $a_i^{\dagger 2} a_i^2$, the Heisenberg equation of motion for any ladder operator (and hence any quadrature) depends on higher-order moments such as $\langle a^\dagger a a \rangle$. This dependence prevents the closure of the moment hierarchy, making it impossible to derive a closed expression for their evolution [26].

Truncating the Heisenberg equations to first-order moments yields a time-dependent shift in the oscillation frequency proportional to the number operator, capturing the intensity-dependent phase rotation characteristic of the Kerr effect. In the steady state, these effective oscillation frequencies take the form

$$\Omega_x^{\text{eff}} = \omega_x + \chi_x(\langle n_x \rangle_{ss} - 1) + \chi_{xy}\langle n_y \rangle_{ss}, \quad (36)$$

$$\Omega_y^{\text{eff}} = \omega_y + \chi_y(\langle n_y \rangle_{ss} - 1) + \chi_{xy}\langle n_x \rangle_{ss}, \quad (37)$$

with $\langle n_i \rangle_{ss}$ representing the mean value of the number operator of the i -th mode in the steady state.

Kerr nonlinearities were introduced into the simulations, with strength parameters set to $\chi_x = 2$ and $\chi_y = 3$, while keeping a damping rate $\kappa = 0.1$ and no cross term χ_{xy} . The optimal cutoff rank increased to $r_{\text{opt}} = 20$, five times more than in the previous case. Taking into account that the value of r_{opt} determines the number of main dynamical modes identified, four of those 20 were classified as nonlinear. Their appearance can be regarded as a consequence of the lack of a closed-form expression for the evolution of the quadratures, as previously discussed. Although this represents a theoretical limitation, the algorithm can successfully reconstruct the system dynamics using the forced linear model built on the quadratures. Figures 4(a) and 4(c) illustrate the reconstructed phase space, showing a spiral trajectory that collapses smoothly towards the origin, with each turn occurring at a slightly different instantaneous frequency, consistent with the presence of a Kerr nonlinearity.

Regarding the retrieval of Hamiltonian parameters, the eigenvalues of \mathbf{A} appeared in complex-conjugate pairs, similar to the previous case. Their real parts, associated with damping rates, did not exhibit any identifiable pattern. In contrast, the imaginary part followed the relation $\pm(\hat{\omega}_i + n\hat{\chi}_i)$, where $n \in \mathbb{N}$. Hence, *mHAVOK* was capable of recovering Kerr nonlinearity coefficients alongside the oscillation frequencies. Such a pattern successfully reproduces the characteristic frequency ladder of Kerr-type nonlinear oscillators.

To retrieve the coefficients $\hat{\chi}_i$, the twenty eigenvalues of \mathbf{A} were first sorted by magnitude. For instance, the imaginary parts of the first eight eigenvalues were $\pm\hat{\omega}_y, \pm(\hat{\omega}_y + \hat{\chi}_y), \pm\hat{\omega}_x$, and $\pm(\hat{\omega}_x + \hat{\chi}_x)$. Considering $\omega_y + \chi_y < \omega_x$, the method ordered the imaginary parts starting from the ones associated with the y -mode. Based on this ordering, the Kerr coefficients were estimated by subtracting the absolute values of the imaginary parts of the third and first eigenvalues to obtain χ_y , and similarly, from the seventh and fifth eigenvalues to obtain χ_x . By applying this procedure, the obtained average percent error was $e_{\chi_i \hat{\chi}_i} = 0.04\%$. Additionally, the associated error

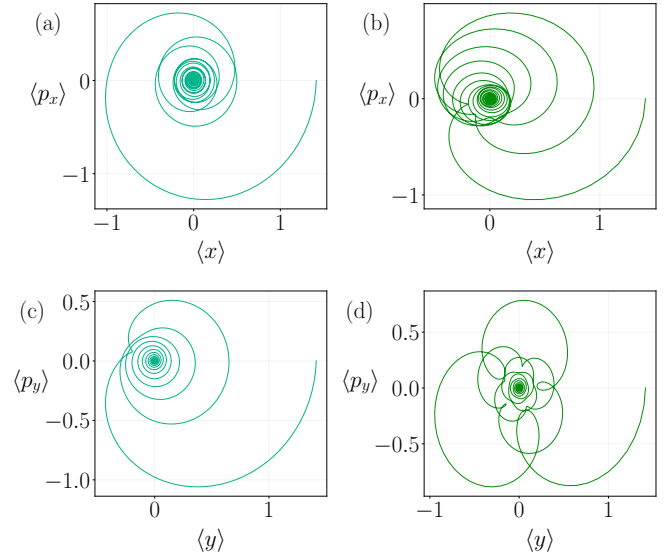


Figure 4. Phase-space reconstruction considering (a-c) $\chi_x, \chi_y = 2, 3$ and (b-d) $\chi_x = \chi_y = 5$. In both cases, the framework successfully recovered Kerr strengths with few numerical errors. No cross-coupling terms were included.

for the oscillation frequencies remained remarkably low, with $e_{\omega_i \hat{\omega}_i} = 0.02\%$.

To observe the evolution of percent error for various Kerr strengths, the simulation was performed over a combination of values ($\chi_x = 2 + 0.3n, \chi_y = 3 + 0.2n$), where $n \in \mathbb{N}$ ranged from 0 to 10. The optimal cutoff rank was calculated for each pair of strengths. Figure 5 shows that the percent error remains below 5% across all cases, even when both χ_x and χ_y are relatively high (greater than 3.0). Similarly, the error in recovering the oscillation frequencies for strong nonlinearities remained less than 5%. In general, the error tended to increase when $\omega_y + \chi_x \approx \omega_x$, as closely spaced frequencies needed to be resolved. However, this increase remained relatively low, with the maximum being 4.96% for the case when $\chi_x = 4.1$ and $\chi_y = 3.2$.

As an example, for the case where both $\chi_x = \chi_y = 5$, i.e., strong nonlinearities, the average percent error for Kerr strength parameters was $e_{\chi_i \hat{\chi}_i} = 1.3\%$, while $e_{\omega_i \hat{\omega}_i} = 1.5\%$. Figures 4(b) and 4(d) show the reconstructed phase space for the above configuration, where now the trajectories form lobes, each corresponding to an instantaneous variation of the phase due to the amplitude-dependent term $\chi_i \langle n_i(t) \rangle$ discussed earlier.

As a final comparison, we examined whether the algorithm could accurately reconstruct the oscillator shown in Figs. 4(b) and 4(d), but under a closed system. Since *mHAVOK* had successfully recovered the quantities of interest for the open system exhibiting strong Kerr nonlinearities, one might expect that reconstructing a closed setup would be straightforward. However, while the reconstruction was successful, as shown in Figs. 6(b) and 6(d), the method proved to be highly sensitive to the

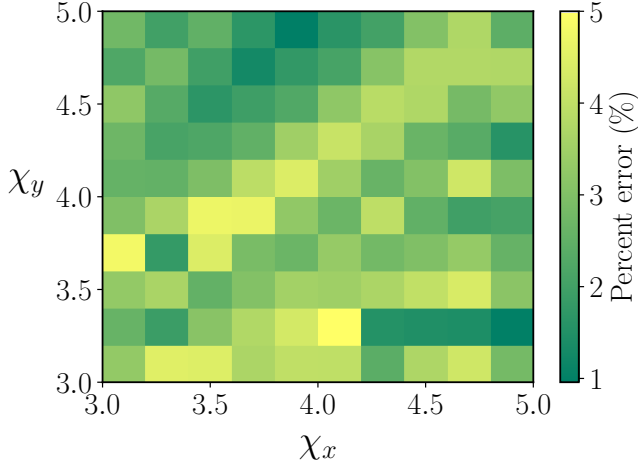


Figure 5. Error comparison for a varying of Kerr nonlinear strengths. The percent error was below 5%, even for strong nonlinearities. To achieve this, the optimal rank r_{opt} was calculated and considered for each simulation.

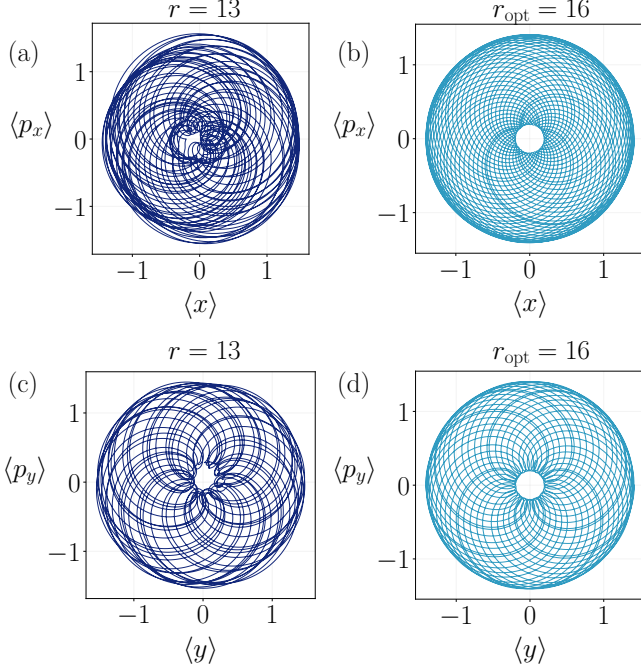


Figure 6. (a-d) Phase-space reconstruction of a closed system considering strong Kerr nonlinearities ($\chi_x = \chi_y = 5$). The dark blue trajectories show the reconstruction using a cutoff rank close to (but not exactly at) the optimal value, while the light blue plots show the reconstruction using the optimal cutoff rank.

cutoff rank r . This contrasts with previous cases, where choosing a rank close to the optimal one generally produced similar results. This is not generally the case, as illustrated in Figs. 6(a) and 6(c).

Table II. Predicted versus actual values of multiple cross-coupling Kerr nonlinearities.

Real value χ_{xy}	Predicted value $\hat{\chi}_{xy}$
0.05	0.07
0.10	0.09
0.15	0.18
0.20	0.19
0.25	0.15

1. Retrieving cross-coupling Kerr terms

We also simulated several cross-term couplings χ_{xy} , while keeping $\chi_x = 0$ and $\chi_y = 0$. In this case, the imaginary part of the eigenvalues of \mathbf{A} is given by $\hat{\Omega}_i = \omega_i + \hat{\chi}_{ij} \langle n_j \rangle$, which is consistent with Eqs. (36) and (37). After retrieving the oscillation frequencies $\hat{\omega}_i$, we determined the value of $\hat{\chi}_{xy}$ using the relationship mentioned above. To achieve this, we tracked the expected value of the number operator for both modes throughout the simulation using QuTiP. After applying this procedure to various values of χ_{xy} , we summarize the corresponding results in Table II. Generally, the method struggles to provide an accurate approximation for moderate and strong cross-coupling terms. This limitation arises because the forced linear model of m HAVOK does not accurately reconstruct the evolution of the number operator. As a result, the oscillation frequencies obtained are only approximated based on the value of the number operator in the steady state.

C. Jaynes-Cummings interaction

The Jaynes-Cummings model describes the interaction between a two-level atom, or qubit, and a single mode of the electromagnetic field [27]. For the following analysis, we assume a coupling on the x -mode of the electromagnetic field. In this situation, the qubit can exist in either the ground state $|g\rangle$ with energy E_g or in the excited state $|e\rangle$ with energy E_e .

The Hamiltonian of the system is $H = H_0 + H_{JC}$, where the Jaynes-Cummings Hamiltonian reads as

$$H_{JC} = H_{\text{atom}} + H_{\text{interaction}}, \quad (38)$$

$$= \frac{\Omega}{2} \sigma_z + g\hbar(a_x \sigma_+ + a_x^\dagger \sigma_-),$$

where $\Omega \equiv E_e - E_g$, g denotes the coupling strength, σ_z is the third Pauli matrix, and (σ_+, σ_-) represent, respectively, the raising and lowering qubit operators. Additionally, the qubit can spontaneously emit a photon and end up in the ground state. This procedure is known as relaxation and is modeled through a jump operator of

the form [28]

$$L^{\text{relaxation}} = \sqrt{\gamma}\sigma_-, \quad (39)$$

where γ is the relaxation rate. Similarly, the process from which the qubit's coherence decays is known as dephasing. If γ_ϕ denotes the dephasing rate, then this process is represented by the jump operator [28]

$$L^{\text{dephasing}} = \sqrt{\gamma_\phi}\sigma_z. \quad (40)$$

In this case, the equation of motion for the annihilation operator is coupled with the qubit lowering operator σ_- . Assuming the atom remains mostly in the ground state allows us to approximate $\sigma_z \approx 1$, leading to the linear set of equations

$$\begin{pmatrix} \dot{a} \\ \dot{\sigma}_- \end{pmatrix} = -i \begin{pmatrix} \omega_x & g \\ -g & \omega_q \end{pmatrix} \begin{pmatrix} a \\ \sigma_- \end{pmatrix}. \quad (41)$$

The eigenvalues of this matrix yield the dressed frequencies that describe the oscillations of the combined qubit-cavity system [27]. Mathematically,

$$\omega_\pm = \frac{1}{2} \left[(\omega_q + \omega_x) \pm \sqrt{4g^2 + \Delta^2} \right], \quad (42)$$

with ω_q being the qubit frequency and $\Delta \equiv \omega_q - \omega_x$ representing the detuning, i.e., how far apart both oscillators are in frequency [28].

Along with the simulation parameters presented in Table 1, we simulated the Jaynes-Cummings interaction with the initial state of the qubit $|g\rangle$ and an oscillation frequency $\omega_q = 2\pi$. The coupling strength was set at $g = 0.15$, while the relaxation and dephasing rates were, respectively, $\gamma = 0.1$ and $\gamma_\phi = 0.01$. For this case, *mHAVOK* was provided with the observables

$$\mathbf{x}' = \begin{bmatrix} \langle x(t) \rangle \\ \langle p_x(t) \rangle \\ \langle y(t) \rangle \\ \langle p_y(t) \rangle \\ \langle \text{Re}(\sigma_+)(t) \rangle \\ \langle \text{Im}(\sigma_+)(t) \rangle \end{bmatrix}. \quad (43)$$

Figure 7(a) shows how the forced linear model successfully reconstructs the parabolic Lissajous curves with decaying amplitudes. In contrast to Fig. 2, this parabola does not collapse; instead, it changes orientation due to the interaction between the qubit and the field. Furthermore, the evolution of the raising qubit operator is also recovered and visualized in Fig. 7(b). Here, coherent Rabi oscillations are prominent in the early times, while the long-term behavior indicates a decay, which can be attributed to the gradual loss of energy and coherence through the jump operators. The optimal cutoff rank was determined to be $r = 20$, with all identified components classified as linear.

Moreover, the imaginary part of the eigenvalues of \mathbf{A} enables the retrieval of the dressed frequencies defined in

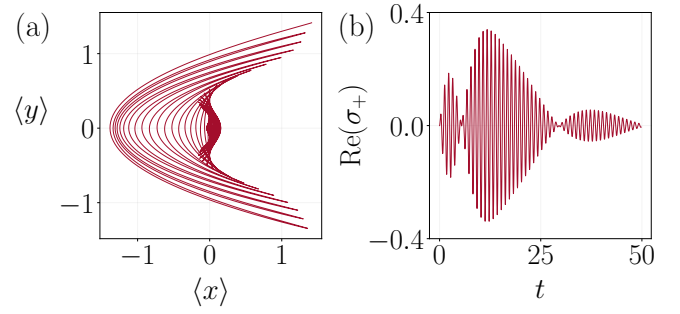


Figure 7. (a) Phase-space reconstruction under a Jaynes-Cummings interaction. (b) Retrieved evolution of $\langle \sigma_+(t) \rangle$ via *mHAVOK*'s forced linear model.

Eq. (42). When solving for $g = 0.15$ in this equation, a retrieved value of $\hat{g} = 0.13$ was obtained, resulting in a percent error of $e_{g\hat{g}} = 13.33\%$. These results were sensitive to moderate to strong dephasing and relaxation rates. However, they improved substantially when the qubit frequency was not a multiple or exactly equal to ω_x . For example, choosing $\omega_q = \pi/e$ resulted in an estimate of $\hat{g} = 0.16$, which has a percent error $e_{g\hat{g}} = 6.67\%$, half of the obtained when $\omega_q = 2\pi$. This happens because when both frequencies coincide, the system is in resonance, making it more difficult to identify the coupling strength reliably.

Varying over multiple values of g , the method accurately recovered only weak and moderate coupling strengths. For example, for $g = 0.25$, the method recovered $\hat{g} = 0.28$, yielding a percent error $e_{g\hat{g}} = 12.00\%$. However, when doubling the coupling strength to $g = 0.50$, the method yielded a percent error of 26%. Therefore, the algorithm provides a reasonable estimation of the coupling parameter when the qubit frequency is not close to the oscillation frequency of the field, and the coupling strength remains within the weak-to-moderate regime.

D. Parametric frequency modulation

The Hamiltonians analyzed so far have been time-independent. In this section, we demonstrate that *mHAVOK* accurately recovers parameters in open systems with time-dependent Hamiltonians. Let us consider the Hamiltonian

$$H = H_0 + H_f(t), \quad (44)$$

where the 2D QHO Hamiltonian H_0 [Eq. (32)] is modified by a time-dependent Hamiltonian with a modulation frequency on the x -mode of the form

$$H_f(t) = \delta \cos(\omega_f t) a_x^\dagger a_x, \quad (45)$$

where δ denotes the modulated amplitude and ω_f the modulated frequency.

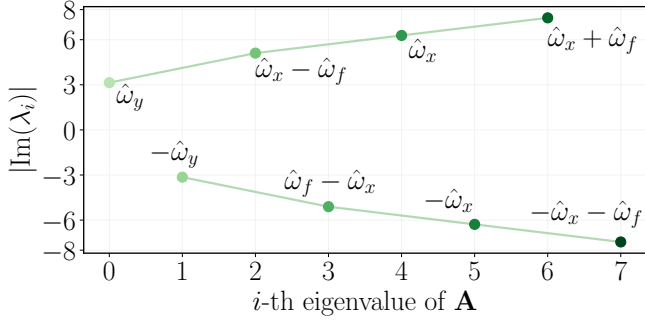


Figure 8. Evolution of the imaginary part of the eigenvalues of \mathbf{A} , which are denoted as λ_i . Complex-conjugate pairs are obtained, each representing a linear combination of the oscillation frequencies ω_x , ω_y and ω_f .

For the Hamiltonian in Eq. (44), the ladder operator and, hence, the quadratures evolve as

$$\langle a_x(t) \rangle = \langle a_x(0) \rangle e^{-\kappa t/2} \sum_{m=-\infty}^{\infty} J_m\left(\frac{\delta}{\omega_f}\right) e^{-i(\omega_x + m\omega_f)t}. \quad (46)$$

Equation (46) exhibits sidebands at $\omega_x + m\omega_f$ weighted by the Bessel coefficients $J_m(\delta/\omega_f)$. In the presence of Kerr and cross-Kerr terms, these sidebands are shifted by the effective frequencies Ω_i^{eff} , combining the effects of frequency and intensity-dependent phase rotations.

By simulating with a modulated amplitude $\delta = 4$ and frequency $\omega_f = \pi/e$, we verified that the algorithm successfully recovered $\hat{\omega}_f$ within 1.74% of their actual value. Figure 8 shows how the imaginary part of the spectrum grows and recovers all oscillation frequencies, even when time-dependent Hamiltonians are present. These can be identified as the expected sidebands presented in Eq. (46). Moreover, as detailed in Sect. II, their appearance can be interpreted as eigenvalues of the Koopman generator restricted to the space of quantum observables, and whose realization is given by \mathbf{A} . This allows us to span a Koopman-invariant subspace from the observables, providing a well-behaved linear evolution of the dynamics, even in the presence of time-dependent parameters. To achieve this, all relevant frequencies must be captured, which is indeed the case.

As a limitation, the method had difficulty retrieving ω_f for large modulation coefficients ($\delta > 3$). For instance, when the modulation amplitude was increased to $\delta = 10$, the spectrum did not show a recognizable pattern. One of the retrieved frequencies was identified as $\hat{\omega}_y$ with an error of 0.06%. However, the preceding and following frequencies could not be expressed as a linear combination of $\hat{\omega}_x$ or $\hat{\omega}_f$. As δ increases, it brings in contributions from more Bessel functions in Eq. (46), leading to the appearance of multiple neighboring and closely spaced sidebands. Therefore, the algorithm is unable to isolate and retrieve the modulated frequency $\hat{\omega}_f$. However, the forced linear model can still reconstruct the system dynamics, mainly because the additional nonlinear

dimensions captured in \mathbf{B} are related to multiple Bessel functions. Despite this capability, the model does not accurately retrieve the modulated frequency.

V. CONCLUSIONS

We demonstrated that the *mHAVOK* method provides an efficient spectral framework for recovering the Hamiltonian parameters of open quantum systems directly from first-moment observables. By connecting the spectrum of the Koopman generator with the eigenvalues of the *mHAVOK* matrix \mathbf{A} , we established a formal bridge between data-driven linear models and the underlying physical dynamics. This connection enables the discrete spectrum of \mathbf{A} to serve as a reliable estimator of the relevant parameters of the system without requiring an analytic solution of the corresponding master equation.

By constructing a forced linear model for the expected value of first-moment observables, spectral analysis accurately retrieves oscillation frequencies, damping rates, and Kerr nonlinearities, even while considering time-dependent Hamiltonians. A reasonable, though not ideal, estimation is achieved for recovering the coupling strengths of a Jaynes-Cummings interaction. Most recovered parameters differ from their actual values by less than 5%, while one of the Jaynes-Cummings coupling strengths shows a maximum deviation of 15%. In these cases, the *mHAVOK* framework identified more nonlinear components, thereby increasing its influence on the system dynamics. Considering the spectral analysis is performed on the matrix that effectively captures the evolution of the linear components, a slight decrease in parameter retrieval accuracy is anticipated. Additionally, the matrix describing the nonlinear components is rectangular in the general case, which prevents performing a spectral analysis.

An advantage of the *mHAVOK* algorithm is its ability to recover strong damping rates, even in the presence of nonlinearities. This value was increased by almost an order of magnitude and was still accurately recovered, at least in the case where neither Kerr nonlinearities nor Jaynes-Cummings interactions were introduced. Furthermore, the method was compared with the FFT and matrix-pencil techniques for recovering oscillation frequencies and damping rates. It outperformed both techniques in strongly damped cases and yielded results equivalent to those of the other methods in other regimes.

Finally, we remark that the method relies on Koopman operator theory to reconstruct the system dynamics. One of the dynamical matrices represents a finite-dimensional realization of the Koopman operator on a Koopman-invariant subspace spanned by the expected values of the observables. To the best knowledge of the authors, this theoretical connection has not been previously reported. Therefore, such results motivate further exploration of Koopman theory as a practical framework

for characterizing quantum dynamical systems.

ACKNOWLEDGMENTS

We acknowledge fruitful discussions with Dr. Oliver Probst's research and running group.

Appendix A: Overview of the *m*HAVOK method

The HAVOK algorithm was originally introduced as a method for studying classical nonlinear dynamical systems via forced linear models [12]. The methodology was mainly applied to attractors and used Takens' embedding theorem, which guaranties a diffeomorphic attractor from delay embeddings of a function of the system's state variables, i.e., an observable [29]. Although HAVOK successfully reconstructs several dynamical systems, multiple limitations have been identified [30, 31]: first, only one function of the system's state variables could be provided as input. Second, out of the r main dynamical modes identified in the schema, it was arbitrarily assumed the system's forcing only came from the r -th mode. Third, no objective criterion for selecting such r was provided.

The limitations mentioned above were recently addressed in a new methodology named *m*HAVOK [11]. The proposed extension allowed the method to include multiple time measurements, included a framework for identifying linear and forcing terms among the main r dynamical modes, and presented an objective criterion for selecting r . Such improvements remarkably improved the reconstructed dynamics, even in complex nonlinear systems. Here we present an explanation of how *m*HAVOK works; for a detailed discussion, we refer the reader to Ref. [11].

1. Review of the *m*HAVOK algorithm.

Consider the evolution of the expected values of observables of a (possibly nonlinear) dynamical system. In this work, we study the expected values of n quantum observables, each represented by a N -dimensional vector and stacked in a vector \mathbf{z}_k ,

$$\mathbf{z}_k := \begin{bmatrix} \langle \hat{O}_0(t_k) \rangle \\ \vdots \\ \langle \hat{O}_{n-1}(t_k) \rangle \end{bmatrix}, \quad (\text{A1})$$

where $k \in \mathbb{N}$ ranges from 0 to $N - 1$. The method performs a delay embedding on \mathbf{z}_k as follows:

$$\mathbf{h}_k := [\mathbf{z}_k, \mathbf{z}_{k+\tau_s}, \dots, \mathbf{z}_{k+(m-1)\tau_s}]^T, \quad (\text{A2})$$

where m is the embedding dimension and $\tau_s = 1$ is known as the delay time. All \mathbf{h}_k are stored in a block Hankel matrix,

$$\mathbf{H} = [\mathbf{h}_0 \quad \mathbf{h}_1 \quad \dots \quad \mathbf{h}_{N-m}] \quad (\text{A3})$$

Then, a Singular Value Decomposition (SVD) is performed on \mathbf{H} ,

$$\mathbf{H} = \mathbf{U}\mathbf{S}\mathbf{V}^T, \quad (\text{A4})$$

where the first r columns of \mathbf{V}^T are retained; let \mathbf{V} denote this new matrix and r be referred to as the cutoff rank. With this considered, a first regression procedure is performed on \mathbf{V} in order to obtain \mathbf{C} , which is defined as the matrix that minimizes

$$\min_{\mathbf{C}} \|\dot{\mathbf{V}} - \mathbf{V}\mathbf{C}^T\|_F, \quad (\text{A5})$$

with $\|\cdot\|_F$ denoting the Frobenius norm and $\dot{\mathbf{V}}$ the temporal derivative of \mathbf{V} . Afterward, the coefficient of determination between each of the columns of $\dot{\mathbf{V}}$ and $\mathbf{V}\mathbf{C}$ is performed,

$$R_i^2 = 1 - \frac{\|\dot{\mathbf{v}}_i - \mathbf{V}\mathbf{c}_i\|^2}{\|\dot{\mathbf{v}}_i - \bar{\dot{\mathbf{v}}}_i \cdot \mathbf{1}\|^2}, \quad (\text{A6})$$

where $(\mathbf{c}_i, \dot{\mathbf{v}}_i)$ are, respectively, the i -th columns of \mathbf{C} and $\dot{\mathbf{V}}$, $\|\cdot\|$ denotes the Euclidean norm in L^2 , $\bar{\dot{\mathbf{v}}}_i$ is the temporal mean of $\dot{\mathbf{v}}_i$ and $\mathbf{1}$ is a vector of ones. Each of the r columns of \mathbf{V} is classified as linear or nonlinear according to their R^2 value:

$$\mathcal{I}_c := \{i : R_i^2 \geq \tau\}, \quad (\text{A7})$$

$$\mathcal{I}_f := \{1, \dots, r\} \setminus \mathcal{I}_c, \quad (\text{A8})$$

with τ being a user-defined threshold and $(\mathcal{I}_c, \mathcal{I}_f)$ representing, respectively, the index sets of linear and nonlinear dimensions. If \mathbf{v}_i is the i -th column of \mathbf{V} , then such components are separated into two different matrices:

$$\mathbf{V}_c := [\mathbf{v}_i]_{i \in \mathcal{I}_c}, \quad \mathbf{V}_f := [\mathbf{v}_i]_{i \in \mathcal{I}_f}. \quad (\text{A9})$$

Next, a second regression problem is defined in order to obtain two new matrices; each of them models, respectively, the evolution of linear (\mathbf{A}) and nonlinear (\mathbf{B}) components. Mathematically,

$$\min_{\mathbf{A}, \mathbf{B}} \|\dot{\mathbf{V}}_c - \mathbf{V}_c \mathbf{A}^T - \mathbf{V}_f \mathbf{B}^T\|_F, \quad (\text{A10})$$

where $\mathbf{A} \in \mathbb{C}^{|\mathcal{I}_c| \times |\mathcal{I}_c|}$ and $\mathbf{B} \in \mathbb{C}^{|\mathcal{I}_c| \times |\mathcal{I}_f|}$ and $|\cdot|$ denoting the cardinality of the sets. Subsequently, the following forced linear system is numerically solved:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \quad (\text{A11})$$

with $\mathbf{u} := \mathbf{V}_f$ and \mathbf{x} representing the simulated state vector. The initial condition $\mathbf{x}(0)$ is a column vector given by the first row of \mathbf{V}_c . Finally, the simulated block Hankel

matrix $\hat{\mathbf{H}}$ is retrieved by reapplying the SVD factorization:

$$\hat{\mathbf{H}} = \mathbf{U}_c \mathbf{\Sigma}_c \mathbf{x}^T, \quad (\text{A12})$$

where \mathbf{U}_c and $\mathbf{\Sigma}_c$ represent the SVD matrices of the linear components. The reconstructed observables defined in Eq. (A1) will be given by the first block of $\hat{\mathbf{H}}$. With this procedure, *m*HAVOK provides a forced linear model that allows for obtaining such quantities of interest.

2. On the identification of the optimal cutoff rank

The optimal cutoff rank r_{opt} is identified by calculating the condition number of the dynamical matrix \mathbf{B} [11], which is defined as the quotient between its highest and lowest singular values [32]. Mathematically,

$$\kappa(\mathbf{B}) := \frac{\sigma_{\max}(\mathbf{B})}{\sigma_{\min}(\mathbf{B})}, \quad (\text{A13})$$

with $\sigma_{\max}(\cdot)$ and $\sigma_{\min}(\cdot)$ each representing the maximum and minimum singular values of \mathbf{B} . Physically, this value represents an optimal, non-redundant choice of dynamical modes from which the system can successfully build a forced linear model of the system dynamics. In general, exceeding such value may lead to overfitting, while not reaching it may prevent crucial dynamical modes from being perceived, precluding a successful reconstruction from being obtained. The condition number, being an indicator of how sensible a matrix is to changes in its input [32], successfully captures these phenomena, therefore being an objective criterion for the cutoff rank selection.

Considering each choice of r has an associated \mathbf{B} , the condition number was calculated over a sweep of values $[r_{\min}, r_{\max}]$, with r_{\min} being the minimum number from which *m*HAVOK identified at least one linear dimension and $r_{\max} = 30$ an arbitrary upper value. The optimal rank employed throughout this work was the one that had the highest condition number within the selected interval.

3. Computational requirements demanded by *m*HAVOK

Considering an input of C vector-valued of observables, each of length N , an embedding dimension m , and a cutoff rank r in the SVD procedure, the computational cost of *m*HAVOK is

$$O[mC(N - m)] + O[mC(N - m) \min(mC, N - m)] + O[(N - m)r^2 + r^3]. \quad (\text{A14})$$

The first term corresponds to the construction of the Block Hankel matrix, the second to its SVD factorization, and the third to the regression procedures and the state-space simulation of the forced linear model. In practice,

$N \gg mC$ and the SVD term dominates. Hence, the time complexity reduces to $O[(mC)^2 N]$, i.e., it increases quadratically with the embedding dimension and linearly with the length of each vector-valued observable.

-
- [1] D. Sank, A. Opremcak, A. Bengtsson, M. Khezri, Z. Chen, O. Naaman, and A. Korotkov, System characterization of dispersive readout in superconducting qubits, *Phys. Rev. Appl.* **23**, 024055 (2025).
 - [2] P. Binder and D. Braun, Quantum parameter estimation of the frequency and damping of a harmonic oscillator, *Phys. Rev. A* **102**, 012223 (2020).
 - [3] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit quantum electrodynamics, *Rev. Mod. Phys.* **93**, 025005 (2021).
 - [4] H. Zhang, B. Pokharel, E. Levenson-Falk, and D. Lidar, Predicting non-Markovian superconducting-qubit dynamics from tomographic reconstruction, *Phys. Rev. Appl.* **17**, 054018 (2022).
 - [5] H. Kamakari, S.-N. Sun, M. Motta, and A. J. Minnich, Digital quantum simulation of open quantum systems using quantum imaginary-time evolution, *PRX Quantum* **3**, 010320 (2022).
 - [6] J. H. M. Jensen and K. Mølmer, Gaussian theory for estimating fluctuating perturbations with backaction-evading oscillator variables, *Phys. Rev. A* **106**, 022613 (2022).
 - [7] T. K. Sarkar and O. Pereira, Using the matrix pencil method to estimate the parameters of a sum of complex exponentials, *IEEE Antennas Propag. Mag.* **37**, 48 (1995).
 - [8] Y. Hua and T. K. Sarkar, Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise, *IEEE Trans. Acoust. Speech Signal Process.* **38**, 814 (1990).
 - [9] C. Pira, Lirandë; Ferrie, On the interpretability of quantum neural networks, *Quantum Machine Intelligence* **6**, 52 (2024).
 - [10] I. A. Luchnikov, E. O. Kiktenko, M. A. Gavreev, H. Ouerdane, S. N. Filippov, and A. K. Fedorov, Probing non-markovian quantum dynamics with data-driven analysis: Beyond “black-box” machine-learning models, *Phys. Rev. Res.* **4**, 043002 (2022).
 - [11] C. Colchero, J. Perez, A. Herrera, and O. Probst, A multichannel generalization of the HAVOK method for the analysis of nonlinear dynamical systems (2025), submitted to *Chaos*, arXiv:2509.13424 [nlin.CD].
 - [12] S. L. Brunton, B. W. Brunton, J. L. Proctor, E. Kaiser, and J. N. Kutz, Chaos as an intermittently forced linear system, *Nat. Commun.* **8**, 19 (2017).
 - [13] A. I. Lvovsky and M. G. Raymer, Continuous-variable optical quantum-state tomography, *Rev. Mod. Phys.* **81**, 299 (2009).
 - [14] H. Arbabi and I. Mezić, Ergodic theory, Dynamic Mode Decomposition, and computation of spectral properties of the Koopman operator, *SIAM J. Appl. Dyn. Syst.* **16**, 2096 (2017).
 - [15] R. Zwanzig, *Nonequilibrium Statistical Mechanics* (Oxford University Press, 2001).
 - [16] Y. T. Lin, Y. Tian, D. Livescu, and M. Anghel, Data-driven learning for the Mori–Zwanzig formalism: A generalization of the Koopman learning framework, *SIAM J. Appl. Dyn. Syst.* **20**, 2558 (2021).
 - [17] I. Mezić, Spectral properties of dynamical systems, model reduction and decompositions, *Nonlinear Dyn.* **41**, 309 (2005).
 - [18] G. Haller and B. Kaszás, Data-driven linearization of dynamical systems, *Nonlinear Dynamics* **112**, 18639 (2024).
 - [19] G. Lindblad, On the generators of quantum dynamical semigroups, *Commun. Math. Phys.* **48**, 119 (1976).
 - [20] D. Manzano, A short introduction to the Lindblad master equation, *AIP Adv.* **10**, 025106 (2020).
 - [21] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, 2007).
 - [22] J. R. Johansson, P. D. Nation, and F. Nori, QuTiP: An open-source Python framework for the dynamics of open quantum systems, *Comput. Phys. Commun.* **183**, 1760 (2012).
 - [23] R. Horn and C. Johnson, *Matrix Analysis* (Cambridge University Press, 2012).
 - [24] D. Walls and G. Milburn, *Quantum Optics* (Springer Berlin Heidelberg, 2008).
 - [25] C. A. Milburn, G. J.; Holmes, Quantum coherence and classical chaos in a pulsed parametric oscillator with a Kerr nonlinearity, *Phys. Rev. A* **44**, 4704 (1991).
 - [26] H. J. Carmichael, *Statistical Methods in Quantum Optics 2: Non-Classical Fields*, 1st ed., Theoretical and Mathematical Physics (Springer Berlin, Heidelberg, 2008).
 - [27] E. T. Jaynes and F. W. Cummings, Comparison of quantum and semiclassical radiation theories with application to the beam maser, *Proc. IEEE* **51**, 89 (1963).
 - [28] J. Larson and T. Mavrogordatos, *The Jaynes–Cummings Model and Its Descendants*, 2053-2563 (IOP Publishing, 2021).
 - [29] F. Takens, Detecting strange attractors in turbulence, in *Dynamical Systems and Turbulence, Warwick 1980*, edited by D. Rand and L.-S. Young (Springer, Berlin, Heidelberg, 1981) pp. 366–381.
 - [30] X. Cai and Y. Qian, Detection and reconstruction of bursting oscillations in complex systems using the HAVOK analysis framework, *Algorithms* **17**, 10.3390/a17090388 (2024).
 - [31] J. Yang, J. Zhao, J. Song, J. Wu, C. Zhao, and H. Leng, A hybrid method using HAVOK analysis and machine learning for predicting chaotic time series, *Entropy* **24**, 10.3390/e24030408 (2022).
 - [32] A. K. Cline, C. B. Moler, G. W. Stewart, and J. H. Wilkinson, An estimate for the condition number of a matrix, *SIAM J. Numer. Anal.* **16**, 368 (1979).