

Data-Driven System Identification of Linear Quantum Systems Coupled to Time-Varying Coherent Inputs

Hendra I. Nurdin, Nina Amini, Jiayin Chen

▶ **To cite this version:**

Hendra I. Nurdin, Nina Amini, Jiayin Chen. Data-Driven System Identification of Linear Quantum Systems Coupled to Time-Varying Coherent Inputs. 2020 59th IEEE Conference on Decision and Control (CDC), Dec 2020, Jeju, South Korea. 10.1109/CDC42340.2020.9303815 . hal-03059873

HAL Id: hal-03059873
https://hal.science/hal-03059873

Submitted on 25 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

arXiv:2012.06040v1 [quant-ph] 11 Dec 2020

Data-Driven System Identification of Linear Quantum Systems Coupled to Time-Varying Coherent Inputs

Hendra I. Nurdin* Nina H. Amini † Jiayin Chen‡

Abstract

In this paper, we develop a system identification algorithm to identify a model for unknown linear quantum systems driven by time-varying coherent states, based on empirical single-shot continuous homodyne measurement data of the system's output. The proposed algorithm identifies a model that satisfies the physical realizability conditions for linear quantum systems, challenging constraints not encountered in classical (non-quantum) linear system identification. Numerical examples on a multiple-input multiple-output optical cavity model are presented to illustrate an application of the identification algorithm.

1 Introduction

Black-box modelling is a modelling paradigm based on learning about a system by observing its response to given inputs, without any prior knowledge of the system's internal structure. It is an important paradigm in science and engineering, in particular in systems and control. For dynamical systems, black-box modelling is achieved through system identification and has a long rich history [1]. In system identification, *single-shot (stochastic) measurement data* (i.e., a single stochastic observation record) collected from a system of interest is recorded against known inputs injected into it and a mathematical model, chosen from a class of models with some unspecified parameters, is fitted based on the data. Stochasticity arises due to internal noise in the system as well as measurement noise.

In the quantum context, parameter estimation and versions of black-box modelling of dynamical quantum systems have been considered in various contexts; see, e.g., [2, 3, 4, 5, 6, 7] and the references therein. Parameter estimation for the class of quantum stochastic input-output models [8, 9, 10, 11], ubiquitous in various physical platforms such as quantum optics, quantum electrodynamical (QED) systems and superconducting circuits, was initiated by Mabuchi [2]. However, the existing methods share one or more of the following features: (i) they were developed for models other than quantum stochastic input-output models (e.g., closed systems with an unknown Hamiltonian) [3, 4, 5, 6, 7], (ii) use repeated projective measurements and averaging rather than a single continuous measurement record [3, 4, 5, 6, 7] or (iii) assume everything is known about the system except for one or a number of unknown parameters [2, 5, 6].

Recent works have investigated fundamental aspects of system identification for quantum input-output systems [12, 13, 14, 15] but no empirical methods have yet been developed for system identification using single-shot continuous measurement data. Such methods are crucial for practical applications of system identification for quantum input-output systems. This paper will begin to close this gap by initiating the study of empirical system identification for the class of linear quantum systems [16, §6.6] [17] based on single-shot continuous measurement data, in the spirit of the classical setting [1]. The possibility of using single-shot measurement data means that quantum input-output systems, such as linear quantum

*H. I. Nurdin is with the School of Electrical Engineering and Telecommunications, UNSW Australia, Sydney NSW 2052, Australia (email: h.nurdin@unsw.edu.au)

†N. H. Amini is a Chargée de Recherche CNRS at Laboratoire des Signaux et Systèmes (L2S), CentraleSupélec, 91190 Gif-sur-Yvette, France (nina.amini@l2s.centralesupelec.fr)

‡J. Chen is with the School of Electrical Engineering and Telecommunications, UNSW Australia, Sydney NSW 2052, Australia (email: jiayin.chen@student.unsw.edu.au)

systems, could potentially be identified much more efficiently compared to other classes of quantum models in term of data collection.

Notation. Throughout the paper, we will use the following notation. $X^\top$ denotes the transpose of a matrix X, $X^\dagger$ denotes the adjoint of a Hilbert space operator X and if $X = [X_{jk}]$ is a matrix of operators then $X^\dagger$ is the conjugate transpose of X, $X^\dagger = [X_{kj}^\dagger]$. I_n will denote an $n \times n$ identity matrix.

2 Linear quantum stochastic systems

Linear quantum stochastic systems, or simply linear quantum systems, are the quantum analogue of linear stochastic systems and represent a collection of quantum harmonic oscillators coupled to one another through a quadratic Hamiltonian as well as being linearly coupled to external bosonic fields. They represent various quantum devices that have linear quantum stochastic evolution in the Heisenberg picture. This includes, for example, optical and superconducting cavities and parametric amplifiers, and gravitational wave interferometers [8, 9, 11]. They are of interest for linear quantum information processing with quantum Gaussian states and gravitational wave interferometry.

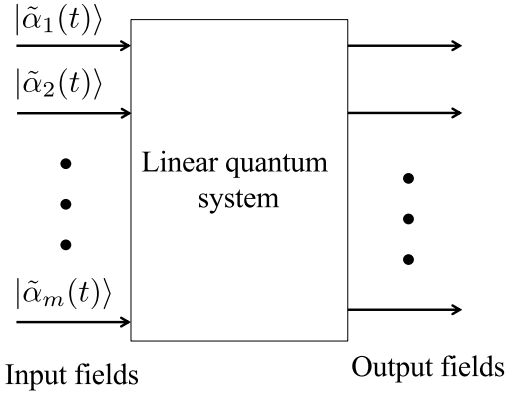


Figure 1: A linear quantum system driven by m fields each in a coherent state

Linear quantum systems are described by a vector $x = (q_1, p_1, q_2, p_2, \ldots, q_n, p_n)^\top$ where q_j and p_j are the position and momentum operators of oscillator j and n is the number of oscillators, a quadratic Hamiltonian $H = \frac{1}{2}x^\top Rx$, where $R = R^\top \in \mathbb{R}^{2n\times 2n}$, a linear coupling operator $L = Kx$ to m external fields with $K \in \mathbb{C}^{m\times 2n}$, and a scattering matrix $S \in \mathbb{C}^{m\times m}$. When the system is driven by m fields that are in a coherent state with amplitude vector $\tilde{\alpha}(t) = \tilde{\alpha}_R(t) + i\tilde{\alpha}_I(t) \in \mathbb{C}^m$, with $\tilde{\alpha}_R(t), \tilde{\alpha}_I(t) \in \mathbb{R}^m$ (see Fig. 1), the joint evolution of the system and field is given by a unitary propagator $U(t)$ solving the Hudson-Parthasarathy quantum stochastic differential equation (QSDE) [18]:

$$dU(t) = \left(-i\left(H + \frac{1}{2}(L+\tilde{\alpha})^\dagger(L+\tilde{\alpha})\right)dt \ + d\mathcal{A}(t)^\dagger(L+\tilde{\alpha}(t)) - (L+\tilde{\alpha}(t))^\dagger d\mathcal{A}(t) \ + \mathrm{Tr}((S-I)d\Lambda(t))\right)U(t),$$

with initial condition $U(0) = I$. In the above QSDE, $\mathcal{A}(t) = [\ \mathcal{A}_1(t) \quad \mathcal{A}_2(t) \quad \ldots \quad \mathcal{A}_m(t)\ ]^\top$ is the vector of annihilation operators for the m field and $\Lambda(t) = [\Lambda_{jk}(t)]_{j,k=1,\ldots,m}$ (with $\Lambda_{jk}^\dagger = \Lambda_{kj}$) satisfying the quantum Itō product rule:

$$d\mathcal{A}_j(t)d\mathcal{A}_k^\dagger(t) = \delta_{jk}dt,\ d\Lambda_{jk}(t)d\Lambda_{uv}(t) = \delta_{ku}d\Lambda_{jv}(t),$$
$$d\Lambda_{jk}(t)d\mathcal{A}_l^\dagger(t) = \delta_{kl}d\mathcal{A}_j^\dagger(t),$$

with all other products between $d\mathcal{A}_j(t)$, $d\mathcal{A}_k^\dagger(t)$ and $d\Lambda_{uv}(t)$ and their adjoints vanishing.

Let $\eta(t) = (\eta_1^q(t), \eta_1^p(t), \eta_2^q(t), \eta_2^p(t), \ldots, \eta_m^q(t), \eta_m^p(t))^\top$ with $\eta_j^q(t) = \mathcal{A}_j(t) + \mathcal{A}_j(t)^\dagger, \eta_j^p(t) = -i\mathcal{A}_j(t) + i\mathcal{A}_j(t)^\dagger$ the amplitude and phase quadratures of the j-th field, respectively. The Heisenberg evolution

$x(t) = U(t)^\dagger x U(t)$ of the vector x of position and momentum operators and the vector of output field $y(t) = U(t)^\dagger \eta(t) U(t)$ are given by the linear QSDE (in the so-called quadrature form [17, Chapter 2]:

$$\begin{aligned} dx(t) &= Ax(t)dt + B(\alpha(t)dt + d\eta(t)) \\ dy(t) &= Cx(t)dt + D(\alpha(t)dt + d\eta(t)). \end{aligned} \tag{1}$$

In the above, $A \in \mathbb{R}^{2n\times 2n}$, $B \in \mathbb{R}^{2n\times 2m}$, $C \in \mathbb{R}^{2m\times 2n}$, $D \in \mathbb{R}^{2m\times 2m}$ and $\alpha(t)$ is $2m \times 1$ vector of real functions representing the phase and amplitude quadratures of coherent amplitudes driving the system, $\alpha(t) = (\tilde{\alpha}_{R,1}(t), \tilde{\alpha}_{I,1}(t), \ldots, \tilde{\alpha}_{R,m}(t), \tilde{\alpha}_{I,m}(t))^\top$, where $\tilde{\alpha}_{s,j}$ is the j-th component of $\tilde{\alpha}_s$, $s \in \{R, I\}$. Similarly, $y(t) = (y_1^q(t), y_1^p(t), y_2^q(t), y_2^p(t), \ldots, y_m^q(t), y_m^p(t))^\top$ is the output field vector containing the amplitude and phase quadratures of the output fields, where y_j^q and y_j^p denote the amplitude and phase quadratures of the j-th field, respectively. Due to quantum constraints, the matrices A, B, C, D need to satisfy the physical realisability constraints [19, 17]:

$$A\mathbb{J}_n + \mathbb{J}_n A^\top + B\mathbb{J}_m B^\top = 0,\ \mathbb{J}_n C^\top + B\mathbb{J}_m D^\top = 0,$$

where $\mathbb{J}_n = I_n \otimes J$ and $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$. If only steady-state measurement data are available, the parameters will only be identifiable up to a similarity transformation, $(A, B, C, D) \to (VAV^{-1}, VB, CV^{-1}, D)$ for some real invertible matrix V [12, 14]. With this transformation, $\mathbb{J}_n$ is replaced with $Z = V\mathbb{J}_n V^\top$ and the physical realizability constraints become:

$$AZ + ZA^\top + B\mathbb{J}_m B^\top = 0 \text{ (I)},\ ZC^\top + B\mathbb{J}_m D^\top = 0 \text{ (II)}. \tag{2}$$

Note that in the above the matrix Z is skew-symmetric, $Z = -Z^\top$ and is required to be invertible.

Information about the system can be obtained by performing measurements on its output. For instance, two basic measurements are $y_q(t) = (y_1^q(t), y_2^q(t), \ldots, y_m^q(t))^\top$ and $y_p(t) = (y_1^p(t), y_2^p(t), \ldots, y_m^p(t))^\top$. These measurements are known as *homodyne measurements* [16, §4.4]. The vector $y_q(t)$ is a homodyne measurement of the amplitude quadrature of the output, while $y_p(t)$ is a homodyne measurement of the phase quadratures. Note that quantum mechanics does not allow simultaneous measurements of $y_q(t)$ and $y_p(t)$ because the elements of these two vectors do not all commute with one another. Thus, it is only meaningful to measure one of these vectors at any time. It follows that,

$$\begin{aligned} dy_q(t) &= C_q x(t)dt + D_q(\alpha(t)dt + d\eta(t)), \\ dy_p(t) &= C_p x(t)dt + D_p(\alpha(t)dt + d\eta(t)), \end{aligned}$$

with $C = \begin{bmatrix} C_q^\top & C_p^\top \end{bmatrix}^\top$ and $D = \begin{bmatrix} D_q^\top & D_p^\top \end{bmatrix}^\top$. It is possible to perform *heterodyne measurement* of y_q and y_p [16, §4.5] which would allow noisy simultaneous measurements of y_p and y_q (but they are not true simultaneous measurements of both quadratures).

When continuous measurement is performed on the quantum system, say by continuously measuring $y_q(t)$, the observed system undergoes a stochastic evolution according to the quantum Kalman filtering equation [17, §4.2]:

$$\begin{aligned} d\hat{x}^q(t) &= A\hat{x}^q(t)dt + B\alpha(t)dt + L_q(t)d\nu_q(t) \\ dy_{qm}(t) &= C_q\hat{x}^q(t)dt + D_q\alpha(t)dt + D_q D_q^\top d\nu_q(t). \end{aligned}$$

Here $y_{qm}(t)$ is the measurement stochastic process (which can be mapped from the operator-valued quantum stochastic process $y_q(t)$ via the Spectral Theorem [20, Theorem 3.3]), $\hat{x}^q$ is the conditional expectation of x^q given the measurement $y_{qm}(t)$ [1] and

$$\nu_q(t) = \left(D_q D_q^\top\right)^{-1}\left(y_{qm}(t) - \int_0^t (C_q\hat{x}^q(\tau) + D_q\alpha(\tau))d\tau\right)$$

[1] $\hat{x}^q$ is also the best mean square estimate of x^q based on $y_{qm}(t)$ [20, 17]

is the so-called *innovation process* of the quantum Kalman filter. Note that $\nu_q(t)$ is a classical standard Wiener process, $\mathbb{E}[\nu_q(t)\nu_q(t')^\top] = \min\{t,t'\}I_n$ that is independent of $\hat{x}^q(s)$ for all $0 \leq s \leq t$. In the quantum Kalman filter, L_q is the Kalman gain and is given by:

$$L_q(t) = Q_q(t)C_q^\top + BD_q^\top,$$

where $Q_q(t) = Q_q(t)^\top \geq 0$ satisfies the Riccati differential equation (RDE):

$$\dot{Q}_q(t) = AQ_q(t) + Q_q(t)A^\top + BB^\top - (Q_q(t)C_q^\top + BD_q^\top)(D_qD_q^\top)^{-1}(Q_q(t)C_q^\top + BD_q^\top)^\top.$$

If the system is asymptotically stable (i.e., the matrix A is Hurwitz), the quantum Kalman filter converges to the steady-state quantum Kalman filter

$$\begin{aligned} d\hat{x}^q(t) &= A\hat{x}^q(t)dt + B\alpha(t)dt + L_q d\nu_q(t) \\ dy_{qm}(t) &= C_q\hat{x}^q(t)dt + D_q\alpha(t)dt + D_qD_q^\top d\nu_q(t). \end{aligned} \tag{3}$$

where L_q is the steady-state Kalman gain given by

$$L_q = Q_qC_q^\top + BD_q^\top, \tag{4}$$

and $Q_q = Q_q^\top \geq 0$ satisfies the algebraic Riccati equation (ARE):

$$AQ_q + Q_qA^\top + BB^\top - (Q_qC_q^\top + BD_q^\top)(D_qD_q^\top)^{-1}(Q_qC_q^\top + BD_q^\top)^\top = 0.$$

Although the equations above are given for measurement of $y_{qm}(t)$, analogous equations can be obtained when measurement of $y_{pm}(t)$ is made.

3 Formulation and numerical solution of identification problem

3.1 Problem formulation

In the system identification problem, we are interested in identifying a model of the form (1) but with system matrices not necessarily of the same dimension, since the true dimensions are not known beforehand, based on the measurement data $y_{qm}(t)$ or $y_{pm}(t)$. In this paper we do not consider heterodyne measurement of y_q and y_p but the approach can be adapted to that case. Throughout, we will consider the system identification problem under the following assumptions:

Assumptions

1. The matrix A is Hurwitz.
2. The data is collected after the system is at steady-state.
3. The matrix D is known. Hence D_q and D_p are known.

An application of standard identification algorithms using knowledge of the single-shot continuous measurement record, say, $y_{qm}(t)$, would identify a model in the innovation form (3) with system matrices $(\hat{A}, \hat{B}, \hat{C}_q, \hat{L}_q)$. However, the identified system matrices from these algorithms will not necessarily satisfy the physical realizability constraints (2) as well as the constraints (4) and (5).

Suppose that we have identified system matrices $(\hat{A}, \hat{B}, \hat{C}_q, \hat{L}_q)$ through some classical identification procedure, such as ARMAX modelling or subspace identification [1, 21, 22]. The remaining problem is to identify system matrices $(\overline{A}, \overline{B}, \overline{C}_q, \overline{L}_q)$ that do satisfy all the constraints required of a linear quantum system. The following standard results will be useful in the ensuing discussion, we include the proofs here for the sake of completeness.

Lemma 1 *Let $\hat{A}$ be Hurwitz. Then the matrix equation $\hat{A}Z + Z\hat{A}^\top = 0$, with Z the same dimension as $\hat{A}$, has the unique solution $Z = 0$.*

Proof. Let z_j denote the j-th column of Z and let $\mathrm{vec}(Z)$ be the vectorization of Z by stacking its columns one on top of the other starting with z_1 at the very top. The equation $\hat{A}Z + Z\hat{A}^\top = 0$ is equivalent to the equation $(\hat{A} \otimes I + I \otimes \hat{A})\mathrm{vec}(Z) = 0$. If $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of A (including their multiplicities), which all have negative real parts, then the eigenvalues of $\hat{A} \otimes I + I \otimes \hat{A}$ are $\lambda_i + \lambda_j$ for $i, j = 1, 2, \ldots, n$. Therefore all eigenvalues of $\hat{A} \otimes I + I \otimes \hat{A}$ also have negative real parts. It follows that the unique solution of $(\hat{A} \otimes I + I \otimes \hat{A})\mathrm{vec}(Z) = 0$ is $\mathrm{vec}(Z) = 0$. Therefore, $Z = 0$ is the unique solution of $\hat{A}Z + Z\hat{A}^\top = 0$. ■

Corollary 2 *Let $\hat{A}$ be Hurwitz. Then the matrix equation $\hat{A}Z + Z\hat{A}^\top + B\mathbb{J}_m B^\top = 0$ has a unique solution Z and this solution is skew-symmetric.*

Proof. Following the proof of Lemma 1, $\hat{A}Z + Z\hat{A}^\top + B\mathbb{J}_m B^\top = 0$ is equivalent to the equation $(\hat{A} \otimes I + I \otimes \hat{A})\mathrm{vec}(Z) = -\mathrm{vec}(B\mathbb{J}_m B^\top)$. By the same argument as in that proof, when $\hat{A}$ is Hurwitz the equation has a unique solution Z, corresponding to $\mathrm{vec}(Z) = -(\hat{A} \otimes I + I \otimes \hat{A})^{-1}\mathrm{vec}(B\mathbb{J}_m B^\top)$. Furthermore, we can also inspect that if Z is a solution then so is $-Z^\top$. Therefore, $Z = -Z^\top$ and the unique solution must be skew-symmetric. ■

In the approach that will be developed below, we first determine $(\overline{A}, \overline{B}, \overline{C}_q)$ (with a Hurwitz $\overline{A}$) and then solve for the Kalman gain $\overline{L}_q$. Given estimates $(\hat{A}, \hat{B}, \hat{C}_q)$, we introduce a loss function $\mathcal{L}$ that is nonnegative function of $\Delta A = \overline{A} - \hat{A}$, $\Delta B = \overline{B} - \hat{B}$ and $\Delta C_q = \overline{C}_q - \hat{C}_q$ with the property that $\mathcal{L}(\Delta A, \Delta B, \Delta C_q) = 0 \Rightarrow \Delta A = 0, \Delta B = 0$ and $\Delta C_q = 0$.

We formulate a linear quantum system identification problem as follows.

Problem 3

$$\underset{\overline{A}, \overline{B}, \overline{C}_q, Z, P}{\text{minimize}} \ \mathcal{L}(\Delta A, \Delta B, \Delta C_q)$$

subject to

$$\begin{aligned}
&P > 0,\\
&P - P^\top = 0,\\
&\overline{A}^\top P + P\overline{A} < 0,\\
&\overline{A}Z + Z\overline{A}^\top + \overline{B}\mathbb{J}_m\overline{B}^\top = 0,\\
&Z\overline{C}_q^\top + \overline{B}\mathbb{J}_m D_q^\top = 0,\\
&Z + Z^\top = 0,\\
&\det(Z)^2 > 0.
\end{aligned} \tag{5}$$

For the loss function $\mathcal{L}$, we choose a simple quadratic function,

$$\begin{aligned}
&\mathcal{L}(\Delta A, \Delta B, \Delta C_q)\\
&= \frac{1}{2}\left(\|\overline{A} - \hat{A}\|_2^2 + \|\overline{B} - \hat{B}\|_2^2 + \|\overline{C}_q - \hat{C}_q\|_2^2\right),
\end{aligned}$$

where $\|X\|_2 = \sqrt{\mathrm{tr}(X^\top X)}$.

After obtaining a solution $(\overline{A}, \overline{B}, \overline{C}_q)$ to the optimization problem 3, we solve for the corresponding Kalman gain $\overline{L}_q$ for the linear quantum system according to (4) and (5).

3.2 Numerical solution

The system identification problem, Problem 3, formulated in the previous section can be viewed as a matrix polynomial programming problem. The objective function is a quadratic function of matrix variables and all the variables are matrix-valued. This is a formidable non-convex optimization problem for which there is no known general solution. Here we borrow a technique proposed in [23] to introduce matrix lifting variables to transform the original matrix polynomial programming problem to a rank

constrained LMI problem. The latter problem can be numerically solved with the LMIRank algorithm [24, 25] (run on the Yalmip toolbox for Matlab [26]) as originally proposed in [23] (see also [17, §5.2.1]).

In the transformation below we will drop the constraint $\det(Z)^2 > 0$ as generically this constraint is expected to be satisfied in the sense that the set where $\det(Z) = 0$ forms a "thin set" in the set of all skew-symmetric matrices in $\mathbb{R}^{2n\times 2n}$; for a discussion of the notion thinness, see, e.g., [27]. To transform the problem we introduce two positive semidefinite symmetric matrix lifting variables $\mathbf{G}_1 \in \mathbb{R}^{(10n+3m)\times(10n+3m)}$ and $\mathbf{G}_2 \in \mathbb{R}^{(4n+2m)\times(4n+2m)}$. We will require these two matrices to satisfy the rank constraints $\mathrm{rank}(\mathbf{G}_1) \leq 2n$ and $\mathrm{rank}(\mathbf{G}_2) \leq 2m$. If these matrices do indeed satisfy the rank constraints then we can factorize them as $\mathbf{G}_j = G_j G_j^\top$ and identify the block elements of G_j as follows:

$$\begin{aligned} G_1^\top &= \begin{bmatrix} I_{2n} & \overline{A}^\top & \overline{A} & \overline{B} & \overline{C}_q^\top & Z^\top & P^\top \end{bmatrix}, \\ G_2^\top &= \begin{bmatrix} I_{2m} & \overline{B}^\top & \mathbb{J}_m^\top \overline{B}^\top \end{bmatrix}. \end{aligned} \tag{6}$$

Now, let $\mathbf{G}_j(k,l)$ denote the (k,l)-th block matrix in $\mathbf{G}_j$. If the $\mathbf{G}_j$ matrices satisfy the specified rank constraints then we have the identification $\mathbf{G}_j(k,l) = G_j(k)G_j(l)^\top$, where $G_j(k)$ denotes the k-th block element of G_j according to the block partitioning in (6). In terms of these block matrices the cost function L can be written as

$$\begin{aligned} \mathcal{L}(\Delta A, \Delta B, \Delta C_q) = \frac{1}{2}\Big(&\mathrm{Tr}[\mathbf{G}_1(2,2) + \hat{A}\hat{A}^\top] + \mathrm{Tr}[\mathbf{G}_1(4,4) + \hat{B}\hat{B}^\top] \\ &+\mathrm{Tr}[\mathbf{G}_1(5,5) + \hat{C}_q\hat{C}_q^\top] - 2\mathrm{Tr}[\hat{A}^\top \mathbf{G}_1(2,1)] \\ &-2\mathrm{Tr}[\hat{B}^\top \mathbf{G}_1(1,4)] - 2\mathrm{Tr}[\hat{C}_q^\top \mathbf{G}_1(5,1)]\Big) \end{aligned}$$

and the constraints (5) can be written as

$$\begin{aligned} &\mathbf{G}_1(1,7) - \mathbf{G}_1(7,1) = 0, \\ &\mathbf{G}_1(1,7) \geq \epsilon I_{2n}, \\ &\mathbf{G}_1(3,7) + \mathbf{G}_1(7,3) \leq -\epsilon \mathbf{G}_1(1,7), \\ &-\mathbf{G}_1(2,6) + \mathbf{G}_1(6,2) + \mathbf{G}_2(3,2) = 0, \\ &\mathbf{G}_1(6,5) + \mathbf{G}_1(1,4)\mathbb{J}_m D_q^\top = 0, \\ &\mathbf{G}_1(1,6) + \mathbf{G}_1(6,1) = 0. \end{aligned}$$

where $\epsilon > 0$ (we set $\epsilon = 10^{-3}$ throughout) and the last constraint ensures the solution for Z returned by the algorithm is skew-symmetric. The constant ϵ has been introduced to replace strict inequality constraints with non-strict ones, as required for the numerical software packages that will be used. From (6), we obtain the following auxiliary constraints on the block elements of $\mathbf{G}_j(k,l)$:

$$\begin{aligned} \mathbf{G}_1(1,1) - I_{2n} &= 0, \\ \mathbf{G}_2(1,1) - I_{2m} &= 0, \\ \mathbf{G}_1(1,3) - \mathbf{G}_1(2,1) &= 0, \\ \mathbf{G}_1(1,4) - \mathbf{G}_2(2,1) &= 0, \\ \mathbf{G}_2(3,1) - \mathbf{G}_2(2,1)\mathbb{J}_m &= 0, \\ \mathbf{G}_i \geq 0, \quad i &= 1,2 \end{aligned}$$

and the original rank constraints

$$\mathrm{rank}(\mathbf{G}_1) \leq 2n, \quad \mathrm{rank}(\mathbf{G}_2) \leq 2m.$$

We remark that if the above constraints are satisfied the original variables of the problem can be recovered from the corresponding block elements of $\mathbf{G}_j$, according to (6). We then solve for the corresponding Kalman gain $\overline{L}_q$ for the identified linear quantum system according to (4) and (5).

To solve this rank-constrained LMI problem, we employ the LMIRank algorithm in [25]. The initial guess for the algorithm is chosen to be $\hat{\mathbf{G}}_j = \hat{G}_j \hat{G}_j^\top$, where $\hat{G}_j^\top$ is obtained from $G_j^\top$ by replacing the

variables $(\overline{A}, \overline{B}, \overline{C}_q)$ with $(\hat{A}, \hat{B}, \hat{C}_q)$. We set the initial guess for P as a solution to the LMI $\hat{A}^\top P + P\hat{A} < 0, P > 0$ and the initial guess for Z to be $\mathbb{J}_n$.

The LMIRank algorithm only solves a feasibility problem. To minimize the cost function, we employ a standard bisection strategy by including $\mathcal{L}(\Delta A, \Delta B, \Delta C_q) \leq \gamma$ as an additional constraint in the feasibility problem. Starting with an initial guess, we half γ each time the LMIRank algorithm returns a feasible solution. Otherwise, we set $\gamma = 1.2\gamma$.

4 Numerical examples

To test the proposed identification method, we will use simulated data of quadrature measurements at the output of a linear quantum system. This can be done in a standard way by generating a sample of a band-limited approximation of the standard white noise vector $\dot{\nu}_q(t)$ (or $\dot{\nu}_p(t)$ depending on the measurement being considered) satisfying $\mathbb{E}[\dot{\nu}_q(t)\dot{\nu}_q(t')^\top] = \delta(t-t')I$, and numerically integrating the SDE for the quantum Kalman filter (3) with a small sampling time of T_s to generate $\dot{y}_{qm}$ (y_{qm} is just the integral of $\dot{y}_{qm}$). We use time derivatives because classical linear system identification algorithms implemented in Matlab use the derivative $\dot{y}_{qm}$ as the input data.

As a numerical example, we consider identifying a multiple-input multiple-output optical cavity with position and momentum operators q and p. Here $H = \Delta(q^2+p^2)/2$ and

$$L = [\ \sqrt{\kappa_1}(q+ip)/2 \quad \sqrt{\kappa_2}(q+ip)/2 \quad \sqrt{\kappa_3}(q+ip)/2\ ]^\top,$$

with $\Delta = 10$, $\kappa_1 = 5$, $\kappa_2 = 3$, and $\kappa_3 = 2$, and $S = I_3$, corresponding to the system matrices,

$$A = \begin{bmatrix} -5 & 20 \\ -20 & -5 \end{bmatrix},\ C = \begin{bmatrix} 2.2361 & 0 \\ 0 & 2.2361 \\ 1.7321 & 0 \\ 0 & 1.7321 \\ 1.4142 & 0 \\ 0 & 1.4142 \end{bmatrix},\ D = I_6.$$

$$B = \begin{bmatrix} -2.2361 & 0 & -1.7321 & 0 & -1.4142 & 0 \\ 0 & -2.2361 & 0 & -1.7321 & 0 & -1.4142 \end{bmatrix},$$

Using a sampling time of $T_s = 10^{-2}$ s, we generate the measurement data from the system ($\dot{y}_{pm}$ and $\dot{y}_{qm}$) for a total time duration of 80 s, with initial state $\hat{x}^j(0) = 0$, where $j \in \{q, p\}$. The first 20 seconds of the data is for driving the system to its steady state and is not used for identification. The next 30 seconds of the data is used for model estimation and the last 30 seconds is for model validation. The system is excited by a pseudo-random binary sequence (PRBS) generated using the "idinput" Matlab command, a persistently exciting input signal [1, Chapter 13]. The amplitudes of the PRBS are set to be $\Omega = \{10/\sqrt{T_s}, 50/\sqrt{T_s}, 100/\sqrt{T_s}\}$ to investigate the effect of different signal-to-noise ratio (SNR) in the presence of white noise on the estimated models. We employ subspace identification [21, 22] through the "n4sid" Matlab command to estimate the system matrices. As the order of estimated models is unknown a priori, classical (non-physically realizable) models of state-space dimension $2n \in \{2, 4, 6\}$ are identified and compared using their "relative energy" contributions, as computed and plotted by the n4sid command. States with small relative energies contribute little to the model accuracy and can be discarded with little impact. Table 1 and Table 2 show the relative energy contributions of estimated classical models using measurement data $\dot{y}_{qm}$ and $\dot{y}_{pm}$, respectively. For all values of Ω, relative energy suggests that the simplest model with $n = 1$ is sufficient. As Ω increases, relative energy for $n = 1$ further increases.

From the classical models produced by the subspace identification, we then identify system matrices $(\overline{A}_j, \overline{B}_j, \overline{C}_j)$ that satisfy all constraints (5) of a linear quantum system using the LMIRank algorithm. We observe that the magnitudes of $\hat{B}_j$ estimated by subspace identification are small while the magnitudes of $\hat{C}_j$ are large. To avoid poor numerical conditioning for the LMIRank algorithm, we perform a similarity transformation with $T = 6\Omega I_{2n}$. This transformation leaves $\hat{A}_j$ unchanged but scales $\hat{B}_j$ by 6Ω and $\hat{C}_j$ by $\frac{1}{6\Omega}$. Using the bisection strategy, LMIRank returns the cost function values γ_j tabulated in Table 1

and Table 2. We compute the Akaike final prediction-error (FPE) as in [1, Chapter 16] for the estimated (physically realizable) quantum models obtained after applying the optimization algorithm in Section 3.2. The FPE is defined by

$$\mathrm{FPE}_j = \det\left(\frac{1}{N}\sum_k e_j(kT_s)e_j(kT_s)^\top\right)\frac{1+d/N}{1-d/N},$$

where the summation is over $e_j(kT_s)$ for the validation data (the last 30 s), N is the number of validation data and d is the number of estimated parameters. The prediction error $e_j(kT_s)$ is obtained using the "resid" Matlab command. We also report the percentage fit for each output, defined by

$$\mathrm{Fit}_{j,l} = \left(1 - \frac{\sqrt{\sum_k e_{j,l}^2(kT_s)}}{\sqrt{\sum_k(\dot{y}_{jm,l}(kT_s) - \mu_{jm,l})^2}}\right)\times 100\%,$$

where $l = 1,\ldots,m$, $e_{j,l}(kT_s)$ and $\dot{y}_{jm,l}(kT_s)$ are the l-th component of $e_j(kT_s)$ and $\dot{y}_{jm}(kT_s)$, and $\mu_{jm,l} = \frac{1}{N}\sum_k \dot{y}_{jm,l}(kT_s)$. The percentage fits are computed using the "compare" Matlab command.

See Table 1 and Table 2 for FPE_j and $\mathrm{Fit}_{j,l}$ for $j \in \{p,q\}$, respectively.

Table 1: Relative energy contributions of estimated classical models, γ_q, FPE_q and $\mathrm{Fit}_{q,l}$ for estimated quantum models according to measurement data $\dot{y}_{qm}$.

Ω	n	Relative energy	γ_q	FPE_q ($\times 10^6$)	$\mathrm{Fit}_{q,1}$ (%)	$\mathrm{Fit}_{q,2}$ (%)	$\mathrm{Fit}_{q,3}$ (%)
$\frac{10}{\sqrt{T_s}}$	1	7.61	0.0094	1.11	59.8	50.1	42.2
	2	5.08	0.65	1.28	59.2	47.8	41.6
	3	5.00	1.56	1.16	59.8	50.0	42.2
$\frac{50}{\sqrt{T_s}}$	1	9.21	0.004	1.15	91.0	88.4	86.3
	2	5.08	0.65	6.24	87.5	86.1	80.3
	3	5.00	1.56	1.43	90.6	88.4	85.7
$\frac{100}{\sqrt{T_s}}$	1	9.91	0.001	1.14	95.5	94.2	93.1
	2	5.08	0.54	1.17	95.5	94.2	93.1
	3	5.01	1.56	2.46	94.6	93.9	92.0

Table 2: Relative energy contributions of estimated classical models, γ_p, FPE_p and $\mathrm{Fit}_{p,l}$ for estimated quantum models according to measurement data $\dot{y}_{pm}$.

Ω	n	Relative energy	γ_p	FPE_p ($\times 10^6$)	$\mathrm{Fit}_{p,1}$ (%)	$\mathrm{Fit}_{p,2}$ (%)	$\mathrm{Fit}_{p,3}$ (%)
$\frac{10}{\sqrt{T_s}}$	1	7.59	0.015	1.11	58.8	49.3	42.9
	2	5.12	0.65	1.12	58.0	48.6	42.7
	3	5.05	3.73	1.13	58.1	48.8	41.7
$\frac{50}{\sqrt{T_s}}$	1	9.20	0.004	1.11	91.0	88.4	86.1
	2	5.12	0.65	3.0	89.2	86.0	83.8
	3	5.05	3.73	9.5	87.5	82.5	80.3
$\frac{100}{\sqrt{T_s}}$	1	9.90	0.001	1.11	95.4	94.2	93.0
	2	5.12	0.54	1.46	95.3	94.0	92.7
	3	5.05	3.73	22.4	91.2	90.8	88.1

For all values of Ω, estimated physically realizable quantum models with $n = 1$ achieve the smallest γ_j and FPE_j, as well as the best percentage fits. As the signal amplitude increases, for $n = 1$, γ_j decreases from around 0.01 to 0.001 and the percentage fits increase from around 50% to over 90%. In fact, when $\Omega = 100/\sqrt{T_s}$, the estimated classical models with system matrices below almost satisfy the physical

realizability constraints:

$$\hat{A}_q = \begin{bmatrix} -5.22 & -20.05 \\ 19.97 & -4.78 \end{bmatrix},\ \hat{C}_q = \begin{bmatrix} 1.20 & 0.20 \\ 0.93 & 0.15 \\ 0.76 & 0.12 \end{bmatrix} \times 10^4,$$
$$\hat{B}_q = \begin{bmatrix} -4.06 & -0.67 & -3.12 & -0.52 & -2.55 & -0.42 \\ -0.61 & 4.04 & -0.49 & 3.12 & -0.37 & 2.59 \end{bmatrix} \times 10^{-4},$$

and

$$\hat{A}_p = \begin{bmatrix} -4.78 & -20.16 \\ 19.87 & -5.24 \end{bmatrix},\ \hat{C}_p = \begin{bmatrix} -1.19 & -0.20 \\ -0.92 & -0.16 \\ -0.76 & -0.13 \end{bmatrix} \times 10^4,$$
$$\hat{B}_p = \begin{bmatrix} -0.67 & 4.06 & -0.53 & 3.13 & -0.44 & 2.54 \\ 4.01 & 0.75 & 3.08 & 0.54 & 2.59 & 0.45 \end{bmatrix} \times 10^{-4}.$$

This suggests that when $\alpha(t)$ has sufficiently large amplitude (corresponding to a large SNR ratio of the input signal to the quantum noise) the classical subspace identification algorithm is able to produce identified classical models that are close to being physically realizable linear quantum models. To obtain the physically realizable system matrices, we decompose Z_j as $Z_j = V_j \mathbb{J}_n V_j^\top$ for $j \in \{q, p\}$, where

$$Z_q = \begin{bmatrix} 0 & -1.193 \\ 1.193 & 0 \end{bmatrix}, Z_p = \begin{bmatrix} 0 & -1.189 \\ 1.189 & 0 \end{bmatrix},$$
$$V_q = \begin{bmatrix} 1.09 & 0 \\ 0 & -1.09 \end{bmatrix}, V_p = \begin{bmatrix} 0 & -1.09 \\ -1.09 & 0 \end{bmatrix}.$$

Then the corresponding physically realizable system matrices are $(\overline{A}_j, \overline{B}_j, \overline{C}_j) = (V_j^{-1}\overline{A}'_j V_j, V_j^{-1}\overline{B}'_j, \overline{C}'_j V_j)$, where $\overline{A}'_j, \overline{B}'_j, \overline{C}'_j$ are solutions returned by the LMIRank algorithm. Based on measurement data $\dot{y}_{qm}$, we obtain

$$\overline{A}_q = \begin{bmatrix} -5.21 & 20.05 \\ -19.97 & -4.77 \end{bmatrix}, \overline{C}_q = \begin{bmatrix} 2.20 & -0.36 \\ 1.70 & -0.28 \\ 1.40 & -0.23 \end{bmatrix},$$
$$\overline{B}_q = \begin{bmatrix} -2.22 & -0.36 & -1.71 & -0.28 & -1.40 & -0.23 \\ 0.34 & -2.20 & 0.27 & -1.70 & 0.20 & -1.40 \end{bmatrix},$$
$$\overline{L}_q = \begin{bmatrix} -2.04 & -0.75 & 0.68 \\ -0.34 & 0.34 & -0.89 \end{bmatrix} \times 10^{-2}.$$

Based on measurement data $\dot{y}_{pm}$, we obtain

$$\overline{A}_p = \begin{bmatrix} -5.23 & 19.88 \\ -20.17 & -4.77 \end{bmatrix},$$
$$\overline{B}_p = \begin{bmatrix} -2.19 & -0.40 & -1.68 & -0.29 & -1.38 & -0.25 \\ 0.37 & -2.23 & 0.28 & -1.71 & 0.24 & -1.42 \end{bmatrix},$$
$$\overline{C}_p = \begin{bmatrix} 0.37 & 2.18 \\ 0.29 & 1.68 \\ 0.24 & 1.38 \end{bmatrix},$$
$$\overline{L}_p = \begin{bmatrix} -0.33 & 1.43 & 0.81 \\ -2.06 & -1.45 & -2.09 \end{bmatrix} \times 10^{-2}.$$

Furthermore, using the "resid" Matlab command, we observe that the residuals $e_j(kT_s)$ of the estimated quantum models are independent of the inputs and the residuals show no autocorrelation (within 99% confidence interval); see Fig. 2 for the residual sample autocorrelation and [1] for further discussions on

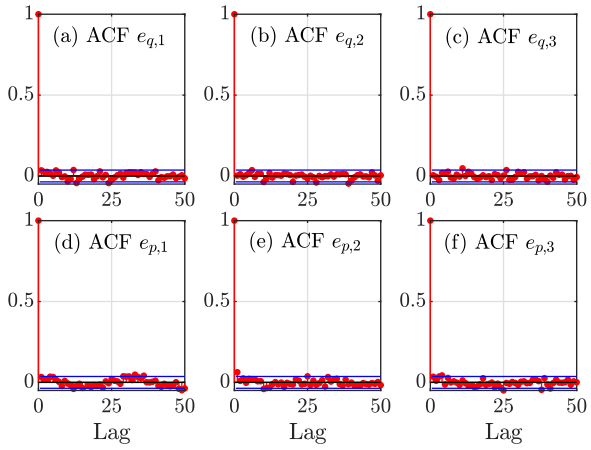


Figure 2: Residual sample autocorrelation of the estimated quantum models for (a) $e_{q,1}(kT_s)$, (b) $e_{q,2}(kT_s)$, (c) $e_{q,3}(kT_s)$, (d) $e_{p,1}(kT_s)$, (e) $e_{p,2}(kT_s)$ and (f) $e_{p,3}(kT_s)$. Horizontal blue lines are the 99% confidence bounds. We show the sample autocorrelation up to 50 lags for illustrative purposes and we observe no sample autocorrelation (within 99% confidence interval) for higher lags.

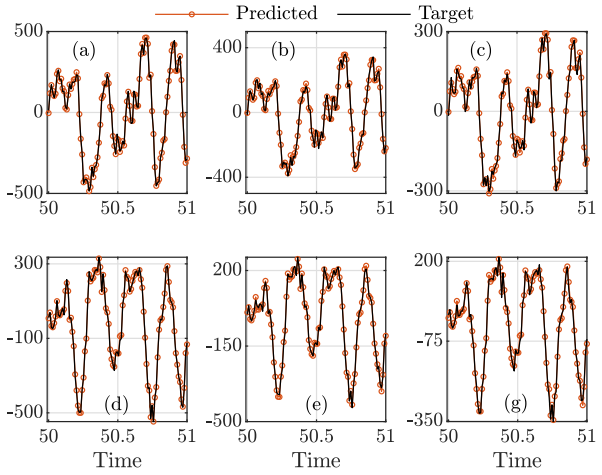


Figure 3: Predicted outputs of the estimated quantum model with $\Omega = 100/\sqrt{T_s}$ and $n = 1$ against target outputs on the first 100 validation data points, for (a) $\hat{\dot{y}}_{qm,1}$ against $\dot{y}_{qm,1}$, (b) $\hat{\dot{y}}_{qm,2}$ against $\dot{y}_{qm,2}$, (c) $\hat{\dot{y}}_{qm,3}$ against $\dot{y}_{qm,3}$, (d) $\hat{\dot{y}}_{pm,1}$ against $\dot{y}_{pm,1}$, (e) $\hat{\dot{y}}_{pm,2}$ against $\dot{y}_{pm,2}$ and (f) $\hat{\dot{y}}_{pm,3}$ against $\dot{y}_{pm,3}$. Here $\hat{\dot{y}}_{jm,l}$ is the l-th component of predicted output $\hat{\dot{y}}_{jm}$ for $j \in \{q, p\}$.

residual diagnostics. Fig. 3 plots the predicted outputs of the quantum model with $\Omega = 100/\sqrt{T_s}$ and $n = 1$ for the first 100 validation data.

We remark that the values of γ_j, FPE_j and $\text{Fit}_{j,l}$ differ for different measurement data $j \in \{q, p\}$. This is due to subspace identification returning different identified system matrices for $(\hat{A}_p, \hat{B}_p)$ and $(\hat{A}_q, \hat{B}_q)$. The two estimates are not expected to be the same as they are estimated using distinct measurement data that are in turn also generated, in general, through distinct stochastic evolutions. It may be possible to develop a technique to merge these two models together to obtain a single identified model but this is beyond the scope of the present work and is a theme for future research.

5 Conclusion

In this paper, based on appropriate assumptions on the system to be identified, we develop a method to identify linear quantum system models based on single-shot continuous stochastic homodyne measurement data generated by the output of unknown linear quantum systems driven by known coherent input fields. The approach involves a two-step procedure. First a (non-physically realizable) classical linear stochastic model is identified using well-established classical system identification algorithms. Then a polynomial matrix feasibility problem is solved to obtain a physically realizable linear quantum system model that is in a sense close to the identified classical stochastic model. We develop a numerical algorithm for solving the polynomial matrix feasibility problem by adopting a matrix lifting technique previously used to numerically solve the coherent quantum LQG problem [23].

We demonstrate our approach in a numerical example. The numerical algorithm is able to identify a multiple-input multiple-output optical cavity based on simulated single-shot homodyne measurement data for varying amplitudes of the coherent input vector $\alpha(t)$. Although classical identification algorithms cannot in general generate physically realizable linear quantum models, our numerical examples indicate that for $\alpha(t)$ with sufficiently high amplitude the classical identified models produced by classical subspace identification can be close to being physically realizable. That is, the identified system matrices almost satisfy the physical realizability constraints of linear quantum systems. However, in practice, high amplitude inputs may not be achievable or consume too much energy to be generated. The case of much practical interest is the one with lower power inputs and this is where the method developed here will be of interest.

The current work assumes the simplification of knowing the output feedthrough matrix D, which in general is not the case. Future work can include generalizing the proposed approach to remove this assumption, developing improved numerical algorithms and finding a method to combine the two models obtained by different measurement quadratures in order to identify a single model.

References

[1] L. Ljung, *System Identification: Theory for the User*, 2nd ed. Prentice-Hall, 1999.

[2] H. Mabuchi, "Dynamical identification of open quantum systems," *Quantum Semiclass. Opt.*, vol. 8, p. 1103, 1996.

[3] D. Burgarth and K. Yuasa, "Quantum system identification," *Phys. Rev. Lett.*, vol. 108, p. 080502, 2012.

[4] J. Zhang and M. Sarovar, "Quantum hamiltonian identification from measurement time traces," *Phys. Rev. Lett.*, vol. 113, p. 080401, 2014.

[5] ——, "Identification of open quantum systems from observable time traces," *Phys. Rev. A*, vol. 91, p. 052121, 2015.

[6] A. Sone and P. Cappellaro, "Hamiltonian identifiability assisted by a single-probe measurement," *Phys. Rev. A*, vol. 95, p. 022335, 2017.

[7] Y. Wang, D. Dong, J. Zhang, I. R. Petersen, and H. Yonezawa, "A quantum Hamiltonian identification algorithm:computational complexity and error analysis," *IEEE Transactions Automat. Control*, vol. 63, no. 5, pp. 1388–1403, 2018.

[8] C. W. Gardiner and P. Zoller, *Quantum Noise: A Handbook of Markovian and Non-Markovian Quantum Stochastic Methods with Applications to Quantum Optics*, 3rd ed. Berlin and New York: Springer-Verlag, 2004.

[9] J. Gough and M. R. James, "The series product and its application to quantum feedforward and feedback networks," *IEEE Trans. Automat. Control*, vol. 54, no. 11, pp. 2530–2544, 2009.

[10] ——, "Quantum feedback networks: Hamiltonian formulation," *Comm. Math. Phys.*, vol. 287, pp. 1109–1132, 2009.

[11] J. Combes, J. Kerckhoff, and M. Sarovar, "The SLH framework for modeling quantum input-output networks," *Adv. Phys. X*, vol. 2, no. 784, 2017.

[12] M. Guţă and N. Yamamoto, "System identification for passive linear quantum systems," *IEEE Trans. Automat. Contr.*, vol. 61, no. 4, pp. 921–936, 2016.

[13] M. Guţă and J. Kiukas, "Information geometry and local asymptotic normality for multi-parameter estimation of quantum markov dynamics," *J. Math. Phys.*, vol. 58, no. 052201, p. 052201, 2017.

[14] M. Levitt and M. Guţă, "Identification of single-input–single-output quantum linear systems," *Phys. Rev. A*, vol. 95, p. 033825, 2017.

[15] M. Levitt, M. Guţă, and H. I. Nurdin, "Power spectrum identification for quantum linear systems," *Automatica*, vol. 90, pp. 255–262, 2018.

[16] H. M. Wiseman and G. J. Milburn, *Quantum Measurement and Control*. Cambridge University Press, 2010.

[17] H. I. Nurdin and N. Yamamoto, *Linear Dynamical Quantum Systems: Analysis, Synthesis, and Control*, ser. Communications and Control Engineering. Cham: Switzerland: Springer, 2017.

[18] K. Parthasarathy, *An Introduction to Quantum Stochastic Calculus*. Berlin: Birkhauser, 1992.

[19] M. R. James, H. I. Nurdin, and I. R. Petersen, "H^∞ control of linear quantum stochastic systems," *IEEE Trans. Automat. Control*, vol. 53, no. 8, pp. 1787–1803, 2008.

[20] L. Bouten, R. van Handel, and M. R. James, "An introduction to quantum filtering," *SIAM J. Control Optim.*, vol. 46, pp. 2199–2241, 2007.

[21] P. van Overschee and B. de Moor, *Subspace Identification for Linear Systems: Theory-Implementation-Applications*. Kluwer Academic Publishers, 1996.

[22] S. J. Qin, "An overview of subspace identification," *Comp. Chem. Eng.*, vol. 30, pp. 1502–1513, 2006.

[23] H. I. Nurdin, M. R. James, and I. R. Petersen, "Coherent quantum LQG control," *Automatica*, vol. 45, pp. 1837–1846, 2009.

[24] R. Orsi, "LMIRank: Software for rank constrained LMI problems," 2005. [Online]. Available: http://rsise.anu.edu.au/~{}robert/lmirank/

[25] R. Orsi, U. Helmke, and J. B. Moore, "A Newton-like method for solving rank constrained linear matrix inequalities," *Automatica*, vol. 42, no. 11, pp. 1875–1882, 2006, extended version available at R. Orsi's homepage.

[26] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004. [Online]. Available: http://control.ee.ethz.ch/~{}joloef/yalmip.php

[27] H. I. Nurdin, S. Grivopoulos, and I. R. Petersen, "The transfer function of generic linear quantum stochastic systems has a pure cascade realization," *Automatica*, vol. 69, pp. 324–333, 2016.