Direct coupling coherent quantum observers with discounted mean square performance criteria and penalized back-action

Igor G. Vladimirov · Ian R. Petersen

Received: 18 June 2021 / Accepted: 5 January 2022 / Published online: 5 March 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

Abstract
This paper is concerned with quantum harmonic oscillators consisting of a quantum plant and a directly coupled coherent quantum observer. We employ discounted quadratic performance criteria in the form of exponentially weighted time averages of the second-order moments of the system variables. Small-gain-theorem bounds are obtained for the back-action of the observer on the covariance dynamics of the plant in terms of the plant–observer coupling. A coherent quantum filtering (CQF) problem is formulated as the minimization of the discounted mean square of an estimation error, with which the dynamic variables of the observer approximate those of the plant. The cost functional also involves a quadratic penalty on the plant–observer coupling matrix in order to mitigate the back-action effect. For the discounted mean square optimal CQF problem with penalized back-action, we establish the first-order necessary conditions of optimality in the form of algebraic matrix equations. By using the Hamiltonian structure of the Heisenberg dynamics and Lie-algebraic techniques, this set of equations is represented in a more explicit form for equally dimensioned plant and observer. For a class of such observers with autonomous estimation error dynamics, we obtain a solution of the CQF problem and outline a homotopy method. The computation of the performance criteria and the observer synthesis are illustrated by numerical examples.

Keywords Quantum harmonic oscillator · Direct coupling · Coherent quantum filtering · Observer back-action · Discounted mean square optimality · Hamiltonian matrix · Lie algebra

This work is supported by the Australian Research Council Grant DP210101938.
1 Introduction

Noncommutative counterparts of classical control and filtering problems [23] are a subject of active research in quantum control which is concerned with dynamical and stochastic systems governed by the laws of quantum mechanics and quantum probability [17, 27]. These developments (see, for example, [21, 28, 30, 47, 48]) are particularly focused on open quantum systems whose internal dynamics are affected by interaction with the environment [6]. In such systems, the evolution of dynamic variables (as noncommutative operators on a Hilbert space) is often modelled using the Hudson–Parthasarathy calculus [19, 31] (see also [16]) which provides a rigorous framework of quantum stochastic differential equations (QSDEs) driven by quantum Wiener processes on symmetric Fock spaces. In particular, linear QSDEs model open quantum harmonic oscillators (OQHOs) [11] whose dynamic variables (such as the position and momentum or annihilation and creation operators [37]) satisfy canonical commutation relations (CCRs). This class of QSDEs is important for linear quantum control theory [33] and applications to quantum optics [13, 53] which provides one of platforms for quantum information technologies [29].

One of the fundamental problems for quantum stochastic systems is the coherent quantum linear quadratic Gaussian (CQLQG) control problem [30] which is a quantum mechanical counterpart of the classical LQG control problem. The latter is well known in linear stochastic control theory due to the separation principle and its links with Kalman filtering and deterministic optimal control settings such as the linear quadratic regulator (LQR) problem [23]. Coherent quantum feedback control [24, 56] employs the idea of control by interconnection, whereby quantum systems interact with each other directly or through optical fields in a measurement-free fashion, which can be described using the quantum feedback network formalism [14]. In comparison with traditional observation-actuation control, coherent quantum control avoids the “lossy” conversion of operator-valued quantum variables into classical signals (which underlies the quantum measurement process), is potentially faster and can be implemented on micro- and nano-scales using natural quantum mechanical effects.

In coherent quantum filtering (CQF) problems [28, 48], which are “feedback-free” versions of the CQLQG control problem, an observer is cascaded in a measurement-free fashion with a quantum plant so as to develop quantum correlations with the latter over the course of time. Both problems employ mean square performance criteria and involve physical realizability (PR) constraints [21, 38] on the state-space matrices of the quantum controllers and filters. The PR constraints are a consequence of the specific Hamiltonian structure of quantum dynamics and complicate the design of optimal coherent quantum controllers and filters. Variational approaches of [46–48] reformulate the underlying problem as a constrained covariance control problem and employ an adaptation of ideas from dynamic programming, the Pontryagin minimum principle [42] and nonlinear functional analysis. In particular, the Frechet differentiation of the LQG cost with respect to the state-space matrices of the controller or...
filter subject to the PR constraints leads to necessary conditions of optimality in the form of nonlinear algebraic matrix equations. Although this approach is quite similar to [2, 41] (with the quantum nature of the problem manifesting itself only through the PR constraints), the resulting equations appear to be much harder to solve than their classical predecessors.

Fully quantum variational techniques, using perturbation analysis [39, 49, 50, 52] beyond the class of OQHOs, and symplectic geometric tools [40], suggest that the complicated sets of nonlinear equations for optimal quantum controllers and filters may appear to be more amenable to solution if they are approached using Hamiltonian structures similar to those in the underlying quantum dynamics. Such structures are particularly transparent in closed QHOs. Indeed, these models of linear quantum systems do not involve external bosonic fields and are technically simpler than the above-mentioned OQHOs (although leave room for modelling the latter, for example, through the Caldeira–Leggett infinite system limit using a bath of harmonic oscillators [7]).

We employ this class of models in the present paper and consider a mean square optimal CQF problem for a plant and a directly coupled observer which form a closed QHO. Since this setting does not use quantum Wiener processes, it simplifies the technical side of the treatment in comparison with [28, 48]. The Hamiltonian of the plant–observer QHO is a quadratic function of the dynamic variables satisfying the CCRs. When the energy matrix, which specifies the quadratic form of the Hamiltonian, is positive semi-definite, the system variables of the QHO are either constant or exhibit oscillatory behaviour. This motivates the use of a cost functional (being minimized) in the form of a discounted mean square of an estimation error (with an exponentially decaying weight [5]) with which the observer variables approximate given linear combinations of the plant variables of interest. The performance criterion also involves a quadratic penalty on the plant–observer coupling in order to achieve a compromise between the conflicting requirements of minimizing the estimation error and reducing the back-action of the observer on the plant. The CQF problem with penalized back-action can also be regarded as a quantum-mechanical counterpart to the classical LQR problem. The use of discounted averages of nonlinear moments of system variables and the presence of optimization make this setting different from the time-averaged approach of [34, 35] to CQF in directly coupled QHOs (see [36] for a quantum-optical implementation of that approach).

Since discounted moments of system variables for QHOs play an important role throughout the paper, we discuss the computation of such moments in the state-space and frequency domains for completeness. Using the ideas of the small-gain theorem (see, for example, [9] and references therein) and linear matrix inequalities, we establish upper bounds for the back-action of the observer on the covariance dynamics of the plant in terms of the plant–observer coupling. This leads to a lower bound for the mean square of the estimation error in terms of its value for uncoupled plant and observer. Similar to the variational approach of [47, 48], we develop the first-order necessary conditions of optimality for the CQF problem being considered. These conditions are organized as a set of two algebraic Lyapunov equations (ALEs) for the controllability and observability Gramians which are coupled through another equation for the Hankelian (the product of the Gramians) of the plant–observer composite.
The Hamiltonian structure of the underlying Heisenberg dynamics allows Lie-algebraic techniques (in particular, the Jacobi identity [10]) to be employed in order to represent this set of equations in terms of the commutators of appropriately transformed Gramians. This leads to a more tractable form of the optimality conditions for equally dimensioned plant and observer. We single out a class of such observers with autonomous estimation error dynamics, for which the CQF problem is amenable to numerical solution through a homotopy method (similar to [26]) over the penalty parameter. We also investigate the asymptotic behaviour of the resulting optimal observers in the weak-coupling limit, and illustrate the performance criteria computation and observer synthesis by numerical examples.

The paper is organized as follows. Section 2 specifies the closed QHOs including its subclass with positive semi-definite energy matrices. Section 3 describes the discounted averaging of moments for system operators in such QHOs in the time and frequency domains and illustrates their computation by a numerical example. Section 4 specifies the direct coupling of quantum plants and coherent quantum observers. Section 5 discusses bounds for the observer back-action on the covariance dynamics of the plant. Section 6 formulates the discounted mean square optimal CQF problem with penalized back-action and discusses coupling-estimation inequalities. Section 7 establishes first-order necessary conditions of optimality for this problem. Section 8 represents the optimality conditions in a Lie-algebraic form. Section 9 provides a suboptimal solution of the CQF problem for a class of observers with autonomous estimation error dynamics and gives a numerical example of observer synthesis. Section 10 makes concluding remarks.

Some of the results of this paper were presented without proofs in its brief conference version [51]. In addition to complete proofs, the new results include Theorem 1 and Lemma 2 on the averaging of nonpolynomial and quadratic functions of quantum variables, Sect. 5 on the observer back-action bounds, Sect. 9 on a solution of the CQF problem in the case of autonomous estimation error dynamics, and numerical Examples 1 and 2 on the performance analysis of QHOs and synthesis of observers.

2 Quantum harmonic oscillators

Consider a QHO [37] with an even number \(n\) of dynamic variables \(X_1(t), \ldots, X_n(t)\) which are time-varying self-adjoint operators on a complex separable Hilbert space \(\mathcal{H}\) satisfying the CCRs

\[
[X(t), X(t)^\dagger] := ([X_j(t), X_k(t)])_{1 \leq j, k \leq n} = 2i\Theta, \quad X(t) := \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix}
\]  

at any time \(t \geq 0\) (the time arguments are often omitted). It is assumed that the CCR matrix \(\Theta \in A_n\) is nonsingular. Here, \(A_n\) is the subspace of real antisymmetric matrices of order \(n\). The entries \(\theta_{jk}\) of \(\Theta\) in (1) represent the scaling operators \(\theta_{jk}\mathcal{I}\), with \(\mathcal{I}\) the identity operator on \(\mathcal{H}\). The transpose \((\cdot)^\dagger\) acts on matrices of operators as if
the latter were scalars, vectors are organized as columns unless indicated otherwise, $[\phi, \psi] := \phi \psi - \psi \phi$ is the commutator of operators, and $i := \sqrt{-1}$ is the imaginary unit. The QHO has a quadratic Hamiltonian

$$H := \frac{1}{2} X^T R X,$$

specified by an energy matrix $R \in \mathbb{S}_n$, with $\mathbb{S}_n$ the subspace of real symmetric matrices of order $n$. In view of (1), (2), the Heisenberg dynamics of the QHO satisfy a linear ODE

$$\dot{X} = i[H, X] = AX,$$

where $\dot{()}$ is the time derivative, and $A \in \mathbb{R}^{n \times n}$ is a matrix of constant coefficients given by

$$A := 2 \Theta R.$$

The first equality in (3) applies to the Hamiltonian (regardless of its particular structure (2)) as $\dot{H} = i[H, H] = 0$, whereby $H$ is constant in time. The solution of the ODE (3) is expressed using the matrix exponential as:

$$X(t) = j_t(X_0) := U(t)^{\dagger} X_0 U(t) = e^{i \text{ad}_H(t)} X_0 = e^{tA} X_0,$$

where $\text{ad}_\alpha := [\alpha, \cdot]$, and the subscript $(\cdot)_0$ indicates the initial values at time $t = 0$, so that $X_0 := X(0)$. The first three equalities in (5) apply to a general Hamiltonian $H$ (not necessarily a quadratic function of $X_0$), and $U(t) := e^{-itH}$ is a time-varying unitary operator on $\mathcal{H}$ (with the adjoint $U(t)^{\dagger} = e^{itH}$), which specifies the flow $j_t$ in (5) acting as a unitary similarity transformation on the system variables. The flow $j_t$ preserves the CCRs (1) which, in view of the relation $[X(t), X(t)^{\dagger}] = e^{tA}[X_0, X_0^{\dagger}] e^{tA} = 2i e^{tA} \Theta e^{tA} = 2i \Theta$, are equivalent to the symplectic property $e^{tA} \Theta e^{tA} = \Theta$ of the matrix $e^{tA}$ for any $t$. The infinitesimal form of this property is $A \Theta + \Theta A^T = 0$, which corresponds to the PR conditions for OQHOs [21, 38]. Its fulfillment is ensured by the Hamiltonian structure $A \in \Theta \mathbb{S}_n$ of $A$ in (4). Similar to classical linear systems, if the initial system variables $X_1(0), \ldots, X_n(0)$ of the QHO are in a Gaussian quantum state [8, 32], they remain so over the course of time due to the deterministic linear dependence of $X(t)$ on $X_0$ in (5).

If the energy matrix in (2) is positive semi-definite, $R \succeq 0$ (and hence, has a square root $\sqrt{R} \succeq 0$), then $A = 2 \Theta \sqrt{R} \sqrt{R}$ is isospectral to the matrix $2 \sqrt{R} \Theta \sqrt{R} \in \mathbb{A}_n$ whose eigenvalues are purely imaginary [18]. In the case $R > 0$, this follows directly from the similarity transformation

$$A = R^{-1/2} (2 \sqrt{R} \Theta \sqrt{R}) \sqrt{R}.$$
(see, for example, [34]), whereby $A$ is diagonalized as:

$$A = iV\Omega W, \quad W := V^{-1}, \quad \Omega := \text{diag}(\omega_k). \quad (7)$$

Here, $W := (w_{jk})_{1 \leq j,k \leq n} \in \mathbb{C}^{n \times n}$ is the inverse matrix for a nonsingular matrix $V := (v_{jk})_{1 \leq j,k \leq n} \in \mathbb{C}^{n \times n}$ whose columns $V_1, \ldots, V_n \in \mathbb{C}^n$ are the eigenvectors of $A$, and $\Omega := \text{diag}(\omega_k)_{1 \leq k \leq n} \in \mathbb{R}^{n \times n}$ is a diagonal matrix of frequencies of the QHO. These frequencies (which should not be confused with the eigenvalues of the Hamiltonian $H$ as an operator on $\mathcal{H}$ describing the energy levels of the QHO [37]) are nonzero and symmetric about the origin, and, without loss of generality, are assumed to be arranged as:

$$\omega_k = -\omega_{k+n/2} > 0, \quad k = 1, \ldots, n. \quad (8)$$

Note that $\sqrt{R}V$ is a unitary matrix whose columns are the eigenvectors of $i\sqrt{R}\Theta \sqrt{R} \in \mathbb{H}_n$ in view of (6); see also the proof of Williamson’s symplectic diagonalization theorem [54, 55] in [10, pp. 244–245]. Here, $\mathbb{H}_n := \mathbb{S}_n + i\mathbb{H}_n$ is the subspace of complex Hermitian matrices of order $n$. Substitution of (7) into (5) leads to

$$X(t) = Ve^{it\Omega}WX_0. \quad (9)$$

Since $e^{it\Omega} = \text{diag}(e^{i\omega_k t})_{1 \leq k \leq n}$, the system variables of the QHO in (9) are linear combinations of their initial values whose coefficients are trigonometric polynomials of time:

$$X_j(t) = \sum_{k,\ell=1}^{n} c_{jk\ell}e^{i\omega_k t}X_\ell(0), \quad j = 1, \ldots, n, \quad (10)$$

where $c_{jk\ell}$ are the entries of the complex rank-one matrices

$$C_k := (c_{jk\ell})_{1 \leq j,k,\ell \leq n} = V_kW_k, \quad c_{jk\ell} := v_{jk}w_{k\ell}, \quad (11)$$

with $W_k$ denoting the $k$th row of $W$. The matrices $C_1, \ldots, C_n$ form a resolution of the identity: $\sum_{k=1}^{n} C_k = VW = I_n$, where $I_n$ is the identity matrix of order $n$. Also,

$$\overline{C_k} = C_{k+n/2}, \quad k = 1, \ldots, n. \quad (12)$$

in accordance with (8), whereby (10) is represented in vector-matrix form as:

$$X(t) = \sum_{k=1}^{n/2} (e^{i\omega_k t}C_k + e^{-i\omega_k t}\overline{C_k})X_0 = 2\sum_{k=1}^{n/2} \text{Re}(e^{i\omega_k t}C_k)X_0, \quad (13)$$
with \( \overline{\cdot} \) the complex conjugate. Therefore, for any positive integer \( d \) and any \( d \)-index \( j := (j_1, \ldots, j_d) \in \{1, \ldots, n\}^d \), the following degree \( d \) monomial of the system variables is also a trigonometric polynomial of time \( t \):

\[
\Xi_j(t) := \prod_{s=1}^{d} X_{j_s}(t) = \sum_{k, \ell \in \{1, \ldots, n\}^d} \prod_{s=1}^{d} c_{j_s,k_s} \ell_s e^{i\omega_{k_s} t} \Xi_{\ell}(0).
\] (14)

Here, \( \prod \) denotes the “rightwards” ordered product of operators (the order of multiplication is essential for non-commutative quantum variables), and the sum is over \( d \)-indices \( k := (k_1, \ldots, k_d), \ell := (\ell_1, \ldots, \ell_d) \in \{1, \ldots, n\}^d \). Note that (10) is a particular case of (14) with \( d = 1 \). The relations (9)–(14) remain valid if \( R \gtrsim 0 \), except that (8) is relaxed to the frequencies \( \omega_1, \ldots, \omega_n/2 \) being nonnegative.

3 Discounted moments of system operators

We will be concerned with moments of time-varying operators \( \sigma(t) \) on the Hilbert space \( \mathcal{H} \), which are functions of the system variables \( X_1(t), \ldots, X_n(t) \) of the QHO specified in Sect. 2. For any \( \tau > 0 \), we define a linear functional \( E_\tau \) which maps such a system operator \( \sigma \) to the weighted time average

\[
E_\tau \sigma := \frac{1}{\tau} \int_0^{+\infty} e^{-t/\tau} E\sigma(t) dt.
\] (15)

Here, \( E\sigma(t) := \text{Tr}(\rho \sigma(t)) \) is the quantum expectation over the underlying quantum state \( \rho \) (a positive semi-definite self-adjoint operator on \( \mathcal{H} \) with unit trace). The weighting function \( \frac{1}{\tau} e^{-t/\tau} \) in (15) is the density of an exponential probability distribution with mean value \( \tau \). Therefore, \( \tau \) plays the role of an effective time horizon (ETH) for averaging \( E\sigma(t) \) over time \( t \geq 0 \). This time average (where the relative importance of the quantity of interest decays exponentially) has the structure of a discounted cost functional in dynamic programming problems [5]. In particular, if \( E\sigma(t) \) is right-continuous at \( t = 0 \), then \( \lim_{\tau \to +0} E_\tau \sigma = E\sigma_0 \) is related to the initial system operator \( \sigma_0 := \sigma(0) \). At the other extreme, the infinite-horizon average is defined by

\[
E_\infty \sigma := \lim_{\tau \to +\infty} E_\tau \sigma = \lim_{\tau \to +\infty} \left( \frac{1}{\tau} \int_0^{\tau} E\sigma(t) dt \right),
\] (16)

provided these limits exist. The second of these equalities, whose right-hand side is the Cesaro mean of \( E\sigma(t) \), as a function of time \( t \geq 0 \), follows from the integral version of the Hardy–Littlewood Tauberian theorem [12]. In particular, (16) implies that \( |E_\infty \sigma| \leq \limsup_{\tau \to +\infty} |E\sigma(t)| \).

If the QHO has a positive semi-definite energy matrix, the coefficients in (13), (14) are either constant or oscillatory, which makes the time averages (15), (16) well-defined for nonlinear functions of the system variables and their moments for any \( \tau > 0 \). A similar property underlies applications of harmonic analysis to the heterodyne
detection of signals. To this end, we will use the characteristic function \( \chi_\tau : \mathbb{R} \to \mathbb{C} \) of the exponential distribution and its pointwise convergence:

\[
\chi_\tau(u) := \frac{1}{\tau} \int_0^{+\infty} e^{-t/\tau} e^{iut} \, dt = \frac{1}{1 - iu\tau} \to \delta_{u0} = \begin{cases} 1 & \text{if } u = 0 \\ 0 & \text{if } u \neq 0 \end{cases}, \quad \text{as } \tau \to +\infty, \tag{17}
\]

where \( \delta_{pq} \) is the Kronecker delta. A combination of (14) with (17) implies that if the initial system variables of the QHO have finite mixed moments \( E_\Xi^\ell(0) \) of order \( d \) for all \( \ell \in \{1, \ldots, n\}^d \), then such moments have the following time-averaged values (15):

\[
E_\tau \Xi_j := \frac{1}{\tau} \int_0^{+\infty} e^{-t/\tau} E_\Xi_j(t) \, dt = \sum_{k \in \{1, \ldots, n\}^d} \chi_\tau \left( \sum_{s=1}^{d} \omega_{k_s} \right) \sum_{\ell \in \{1, \ldots, n\}^d} \prod_{s=1}^{d} c_{js,k_s} \ell_s E_\Xi^\ell(0) \tag{18}
\]

for any \( j \in \{1, \ldots, n\}^d \). Hence, the corresponding infinite-horizon average (16) takes the form:

\[
E_\infty \Xi_j = \sum_{k \in \mathcal{K}_d} \sum_{\ell \in \{1, \ldots, n\}^d} \prod_{s=1}^{d} c_{js,k_s} \ell_s E_\Xi^\ell(0), \tag{19}
\]

where \( \mathcal{K}_d := \{(k_1, \ldots, k_d) \in \{1, \ldots, n\}^d : \sum_{s=1}^{d} \omega_{k_s} = 0\} \) is a subset of \( d \)-indices associated with the frequencies \( \omega_1, \ldots, \omega_n \) of the QHO from (7). For every even \( d \), the set \( \mathcal{K}_d \) is nonempty due to the central symmetry (8) of the frequencies. If the system variables of the QHO are in a Gaussian quantum state, then their higher-order moments in (18), (19) are expressed in terms of the first two moments (with \( d \leq 2 \)) by using the Isserlis–Wick theorem [20, 22].

The linear functional \( E_\tau \) in (18) and its limit \( E_\infty \) in (19) are extendable to polynomials and more general functions \( \sigma := f(X) \) of the system variables, provided \( X_0 \) satisfies appropriate integrability conditions. Such an extension of \( E_\infty \), involving the Cesaro mean, is similar to the argument used in the context of Besicovitch spaces of almost periodic functions [3]. If the system is in an invariant state \( \rho \) (so that \( [H, \rho] = 0 \)), then the quantum expectation \( E_\sigma(t) = \text{Tr}(\rho e^{it\text{rad}H}(\sigma_0)) = \text{Tr}(e^{-it\text{rad}H}(\rho)\sigma_0) = \text{Tr}(\rho\sigma_0) = E_\sigma_0 \) is time-independent for any system operator \( \sigma \) evolved by the flow \( j_t \) in (5). In this case, the time averaging in (15) becomes redundant. The subsequent discussion is concerned with general (not necessarily invariant) quantum states \( \rho \).

The following theorem is, in essence, an adaptation of classical results on the averaging of quasi-periodic motions in Hamiltonian systems, where an important role is played by incommensurability; see, for example, [1, pp. 285–289]. Recall that

1 The Gaussian assumption will not be employed in what follows.
real numbers $\zeta_1, \ldots, \zeta_r$ are said to be incommensurable if their linear combination \( \sum_{k=1}^{r} \lambda_k \zeta_k \) with integer coefficients $\lambda_1, \ldots, \lambda_r \in \mathbb{Z}$ vanishes if and only if $\lambda_1 = \cdots = \lambda_r = 0$. The incommensurability means that $\zeta_1, \ldots, \zeta_r$ are not constrained by a linear relation specified by rational coefficients, which (unlike irrational ones) are exceptional in the sense of forming a denumerable subset of the continuum. Accordingly, the incommensurability is a generic property. In particular, if $\zeta_1, \ldots, \zeta_r$ are generated randomly with an absolutely continuous joint probability distribution, then they are almost surely incommensurable.

**Theorem 1** Suppose the energy matrix in (2) satisfies $R > 0$. Also, let the frequencies $\omega_1, \ldots, \omega_{n/2}$ of the QHO, arranged as in (8), be incommensurable. Furthermore, suppose a function $f : \mathbb{R}^n \to \mathbb{R}$ is extended to quantum variables so that

$$g(\phi) := E_f \left( 2 \sum_{k=1}^{n/2} \text{Re}(e^{i\phi_k} C_k) X_0 \right)$$

depends continuously on $\phi := (\phi_k)_{1 \leq k \leq n/2} \in \mathbb{T}^{n/2}$, where $\mathbb{T}$ is the one-dimensional torus implemented as $[0, 2\pi)$, and the matrices $C_k$ are given by (11). Then, the system operator $f(X)$ has the following infinite-horizon average (16):

$$E^\infty f(X) = (2\pi)^{-n/2} \int_{\mathbb{T}^{n/2}} g(\phi) d\phi.$$  

**Proof** From (13), (20), it follows that $E f(X(t)) = g(t\delta)$ for any $t \geq 0$, where $\delta := (\omega_k)_{1 \leq k \leq n/2}$, and the entries of $t\delta \in \mathbb{R}^{n/2}$ are considered modulo $2\pi$. Since the function $g$, which is $2\pi$-periodic in each of its variables, is assumed to be continuous (and hence, $g$ is bounded due to the compactness of the torus), (21) is established as $E^\infty f(X) = \lim_{t \to +\infty} \left( \frac{1}{t} \int_0^t g(t\delta) dt \right) = (2\pi)^{-n/2} \int_{\mathbb{T}^{n/2}} g(\phi) d\phi$. The last equality is obtained by applying the Weyl equidistribution criterion [4] to the map $t \mapsto t\delta$ considered modulo $2\pi$, whereby its sample distribution $D_t(S) := \frac{1}{t} \mu_1 \{ 0 \leq t \leq \tau : t\delta \in S + 2\pi\mathbb{Z}^{n/2} \}$ converges weakly to the uniform probability distribution $(2\pi)^{-n/2} \mu_{n/2}(S)$ on the torus $\mathbb{T}^{n/2}$, provided the frequencies are incommensurable (here, $\mu_r$ is the $r$-dimensional Lebesgue measure). More precisely, $\lim_{t \to +\infty} D_t(S) = (2\pi)^{-n/2} \mu_{n/2}(S)$ for any Borel set $S \subset \mathbb{T}^{n/2}$ whose boundary $\partial S$ satisfies $\mu_{n/2}(\partial S) = 0$. \hfill \Box

The vectors $\delta$ of commensurable frequencies are contained in a denumerable union $\bigcup_{\lambda \in \mathbb{Z}^{n/2} \setminus \{0\}} \lambda \perp := \{ \delta \in \mathbb{R}^{n/2} : \lambda^T \delta = 0 \}$, which has zero $\mu_{n/2}$-measure; see also [1, p. 290]. Therefore, Theorem 1 applies to the infinite-horizon averaging of nonlinear functions of system variables in QHOs with generic spectra.

Of particular use for our purposes is the state-space computation of the discounted time average (15) of second moments of the system variables for finite values of $\tau$ without employing the imaginarity of the spectrum of $A$ and the frequency incommensurability condition of Theorem 1. To this end, note that at every moment of time, $E(X X^T) \in \mathbb{H}^+_n$ due to the generalized Heisenberg uncertainty principle [17], with $\mathbb{H}^+_n$ the set of complex positive semi-definite Hermitian matrices of order $n$. Furthermore,
\[ \text{Im}(XX^T) = \Theta \] remains unchanged in view of the preservation of the CCRs (1). Also, with any Hurwitz matrix \( \alpha \), we associate a linear operator \( L(\alpha, \cdot) \) which maps an appropriately dimensioned matrix \( \beta \) to a unique solution \( \gamma = L(\alpha, \beta) \) of the ALE
\[ \alpha \gamma + \gamma \alpha^T + \beta = 0: \]
\[ L(\alpha, \beta) := \int_{0}^{+\infty} e^{t\alpha} \beta e^{t\alpha^T} dt. \] (22)

The monotonicity of \( L(\alpha, \cdot) \) (with respect to the partial ordering induced by positive semi-definiteness) implies that
\[ L(\alpha, L(\alpha, \beta)) = L(\alpha, \sqrt{\beta} \beta^{-1/2} L(\alpha, \beta) \beta^{-1/2} \sqrt{\beta}) \]
\[ \preceq r(L(\alpha, \beta^{-1}) L(\alpha, \beta)) \] (23)

for any \( \beta > 0 \), in view of the similarity transformation \( \beta^{-1/2} N \beta^{-1/2} \mapsto N \beta^{-1} \), with \( r(\cdot) \) the spectral radius.

**Lemma 1** Let the initial system variables of the QHO have finite second moments \( (E(X_0^T X_0) < +\infty) \) whose real parts form the matrix
\[ \Sigma := \text{Re}E(X_0^T X_0^T). \] (24)

Also, suppose the ETH \( \tau > 0 \) satisfies
\[ \tau < \frac{1}{2 \max(0, \ln r(e^A))}. \] (25)

Then, the matrix of the real parts of the discounted second moments of the dynamic variables can be computed as:
\[ P := \text{Re}E_\tau(XX^T) = \frac{1}{\tau} L(A_\tau, \Sigma) \] (26)

using (22). That is, \( P \) is a unique solution of the ALE
\[ A_\tau P + PA_\tau^T + \frac{1}{\tau} \Sigma = 0, \] (27)

with the Hurwitz matrix
\[ A_\tau := A - \frac{1}{2\tau} I_n. \] (28)

**Proof** From (5), (24), it follows that \( \text{Re}E(X(t)X(t)^T) = e^{tA} \Sigma e^{tA^T} \) for any \( t \geq 0 \). Hence, the corresponding time average (15) takes the form:
\[ P = \frac{1}{\tau} \int_{0}^{+\infty} e^{-t/\tau} \text{Re} \mathbf{E}(X(t)X(t)^T) \, dt = \frac{1}{\tau} \int_{0}^{+\infty} e^{-t/\tau} \mathbf{e}^{A\tau} \mathbf{\Sigma} \mathbf{e}^{A\tau^T} \, dt \]

\[ = \frac{1}{\tau} \int_{0}^{+\infty} \mathbf{e}^{A\tau} \mathbf{\Sigma} \mathbf{e}^{A\tau^T} \, dt = \frac{1}{\tau} \mathbf{L}(A\tau, \mathbf{\Sigma}), \]

thus establishing (26). Here, the matrix \( A\tau \) in (28) is Hurwitz due to (25). \( \square \)

In view of (27), the matrix \( P \) is the controllability Gramian [23] of the matrix pair \( (A\tau, \sqrt{\tau}^{-1} \mathbf{\Sigma}) \). In contrast to similar ALEs for steady-state covariance matrices in dissipative OQHOs [11] (where the matrix \( A \) itself is Hurwitz), the term \( \frac{1}{\tau} \mathbf{\Sigma} \) in (27) comes from the initial condition (24) instead of the Ito matrix of the quantum Wiener process [17, 19, 31]. Since \( A \) is a Hamiltonian matrix (and hence, its spectrum is symmetric about the imaginary axis), the condition (25) is equivalent to the eigenvalues of \( A \) being in the strip \( \{ z \in \mathbb{C} : |\text{Re} z| < \frac{1}{2\tau} \} \). For any \( \tau > 0 \) satisfying (25), a frequency-domain representation of the matrix \( P \) in (26) is

\[ P = \frac{1}{\tau} \text{Re} \int_{0}^{+\infty} \mathbf{e}^{A\tau} \Gamma \mathbf{e}^{A\tau^T} \, dt \]

\[ = \frac{1}{2\pi \tau} \text{Re} \int_{-\infty}^{+\infty} F\left(\frac{1}{2\tau} + i\omega\right) \Gamma F\left(\frac{1}{2\tau} + i\omega\right)^* \, d\omega \]

\[ = \frac{1}{2\pi \tau} \text{Im} \int_{\text{Re} s = \frac{1}{2\tau}} F(s) \Gamma F(s)^* \, ds, \quad (29) \]

with \((\cdot)^* := ((\cdot))^T\) the complex conjugate transpose. Here,

\[ \Gamma := \mathbf{E}(X_0X_0^T) = \mathbf{\Sigma} + i\mathbf{\Theta} \quad (30) \]

is the matrix of second moments of the initial system variables, and

\[ F(s) := (sI_n - A)^{-1} \quad (31) \]

is the transfer function (analytic over \( s \) satisfying \( \text{Re} s > \ln \text{r}(\mathbf{e}^A) \)) which relates the Laplace transform

\[ \tilde{X}(s) := \int_{0}^{+\infty} e^{-st} X(t) \, dt \quad (32) \]

of the quantum process \( X \) from (5) to its initial value \( X_0 \) as \( \tilde{X}(s) = \int_{0}^{+\infty} e^{-t(sI_n - A)} \, dt \)
\[ X_0 = F(s)X_0. \]

The representation (29) is obtained by applying the Plancherel theorem to the Fourier transform \( \int_{0}^{+\infty} e^{-it\omega} \mathbf{e}^{A\tau} \, d\omega = (i\omega I_n - A\tau)^{-1} = F\left(\frac{1}{2\tau} + i\omega\right) \) in view of (28), (31), or an operator version of the theorem to the inverse Fourier transform \( e^{-\frac{t}{\tau} A\tau} X(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega t} \tilde{X}(\frac{1}{2\tau} + i\omega) \, d\omega \) for \( t \geq 0 \) under the condition (25). If \( R > 0 \) (when \( A \) in (4) has a purely imaginary spectrum and (25) holds for arbitrarily large \( \tau \)), the formal limit of (27), as \( \tau \to +\infty \), is \( \mathbf{A} P + \mathbf{P} \mathbf{A}^T = 0 \), which does not have a unique
solution. This non-uniqueness makes the ALE approach of Lemma 1 inapplicable to computing $E_\infty (XX^T)$. An alternative calculation is provided below.

**Lemma 2** Suppose the energy matrix of the QHO in (2) satisfies $R \succ 0$, and the initial system variables have finite second moments assembled into the matrix $\Gamma$ in (30). Then,

$$E_\tau (XX^T) = V(\Phi_\tau \odot (W\Gamma W^*))V^*$$

$$= \sum_{j,k=1}^{n/2} [V_j \ V_j^\ast] \begin{bmatrix} \chi_\tau (\omega_j - \omega_k) & \chi_\tau (\omega_j + \omega_k) \\ \chi_\tau (-\omega_j - \omega_k) & \chi_\tau (\omega_k - \omega_j) \end{bmatrix} \odot \begin{pmatrix} W_j^\ast \Gamma W_k^T \end{pmatrix} \begin{pmatrix} V_k^* \\ V_k^T \end{pmatrix}$$

(33)

for any $\tau > 0$. Here, $V$ is the matrix from (7), use is made of an auxiliary matrix

$$\Phi_\tau := (\chi_\tau (\omega_j - \omega_k))_{1 \leq j,k \leq n}$$

(34)

associated with the frequencies of the QHO through the function $\chi_\tau$ from (17), $\odot$ is the Hadamard product [18], and $C_k$ are the matrices from (11) satisfying (12) under the convention (8). Furthermore, the infinite-horizon time averages of the second moments are computed as:

$$E_\infty (XX^T) = V(\Phi_\infty \odot (W\Gamma W^*))V^*$$

$$= \sum_{j,k=1}^{n/2} \delta_{\omega_j \omega_k} (C_j \Gamma C_k^\ast + \overline{C_j} \Gamma \overline{C_k}^T)$$

$$= \sum_{j,k=1}^{n/2} \delta_{\omega_j \omega_k} [V_j \ V_j^\ast] \begin{bmatrix} W_j \Gamma & 0 \\ 0 & W_j^\ast \Gamma \end{bmatrix} \begin{pmatrix} V_k^* \\ V_k^T \end{pmatrix},$$

(35)

where use is made of a binary matrix

$$\Phi_\infty := (\delta_{\omega_j \omega_k})_{1 \leq j,k \leq n}.$$ 

(36)

**Proof** Although (33) can be obtained from (18) with $d = 2$, we will provide a direct calculation. In view of self-adjointness of the system variables, (9) and (13) imply that

$$X(t)X(t)^\dag = V e^{it\Omega} WX_0X_0^\dag W^* e^{-it\Omega} V^*$$

$$= V(\Psi(t) \odot (WX_0X_0^\dag W^*))V^*$$

$$= \sum_{j,k=1}^{n/2} (e^{i\omega_j t} C_j + e^{-i\omega_j t} \overline{C_j}) X_0X_0^\dag (e^{-i\omega_k t} C_k^\ast + e^{i\omega_k t} \overline{C_k}^T),$$

(37)
with $(\cdot)^\dagger := ((\cdot)\#)^T$ the transpose of the entry-wise operator adjoint $(\cdot)\#$. Here, the diagonal structure of the matrix $\Omega$ in (7) is used together with a complex Hermitian matrix

$$
\Psi(t) := (e^{i(\omega_j - \omega_k)t})_{1 \leq j,k \leq n}
$$

(38)
of rank one, which encodes the time dependence of $XX^T$. The representation (37) allows the time averaging to be decoupled from the quantum expectation as:

$$
E_\tau(XX^T) := \frac{1}{\tau} \int_0^{+\infty} e^{-t/\tau} E(X(t)X(t)^T)dt
$$

$$
= \sum_{j,k=1}^{n/2} \left( \chi_\tau(\omega_j - \omega_k)C_j \Gamma C_k^* + \chi_\tau(\omega_j + \omega_k)C_j \Gamma C_k^T \right)
$$

$$
+ \chi_\tau(-\omega_j - \omega_k)\overline{C}_j \Gamma \overline{C}_k^* + \chi_\tau(\omega_k - \omega_j)\overline{C}_j \Gamma \overline{C}_k^T \right),
$$

(39)

which leads to (33), (34) in view of (11). Here, $\Gamma$ is given by (30), and the relation $\frac{1}{\tau} \int_0^{+\infty} e^{-t/\tau} \Psi(t)dt = \Phi_\tau$ is obtained by applying (17) entrywise to $\Psi$ in (38). The convergence in (17) implies that the matrices (34), (36) are related by $\lim_{\tau \to +\infty} \Phi_\tau = \Phi_\infty$, and $\lim_{\tau \to +\infty} \chi_\tau(\pm(\omega_j + \omega_k)) = 0$ since $\omega_j + \omega_k > 0$ for all $j,k = 1, \ldots, n$ in view of (8). This leads to (35) in view of (39).

The above proof shows that $E_\tau(XX^T)$ is close to $E_\infty(XX^T)$ if the ETH $\tau$ is large in comparison with

$$
\tau_* := \frac{1}{\min(|\omega_j \pm \omega_k| : 1 \leq j,k \leq n/2\{0\})}.
$$

(40)

If the frequencies $\omega_1, \ldots, \omega_{n/2}$ are pairwise different (that is, $\omega_j \neq \omega_k$ for all $1 \leq j < k \leq \frac{n}{2}$, which is a weaker condition than their incommensurability used in Theorem 1), then $\Phi_\infty$ in (36) becomes the identity matrix and (35) reduces to

$$
E_\infty(XX^T) = \sum_{k=1}^{n/2} \left( C_k \Gamma C_k^* + \overline{C}_k \Gamma \overline{C}_k^T \right)
$$

$$
= \sum_{k=1}^{n/2} \left[ V_k \quad \overline{V}_k \right]
$$

$$
\left[ \begin{array}{c} W_k \Gamma W_k^* \\ 0 \end{array} \right] \left[ \begin{array}{c} 0 \\ \overline{W}_k \Gamma \overline{W}_k^T \end{array} \right] \left[ \begin{array}{c} V_k^* \\ V_k^T \end{array} \right],
$$

(41)

Such energy matrices $R > 0$ form an open subset of $S_n$. The corresponding infinite-horizon average of a quadratic form of $X$ is $E_\infty(X^T \Pi X) = \sum_{k=1}^{n/2} \left( V_k^* \Pi V_k W_k \Gamma W_k^* + V_k^T \Pi \overline{V}_k \overline{W}_k \Gamma \overline{W}_k^T \right)$ for any $\Pi \in S_n$. Lemmas 1 and 2 can be used for computing quadratic cost functionals for QHOs, such as the mean square performance criterion.
Fig. 1 The real parts $P_{jk}$ of the discounted second moments of the system variables against the ETH $\tau$ for the two-mode QHO of Example 1. The dashed lines represent the infinite-horizon averages as $\tau \to +\infty$, and the “◦”s indicate the initial values $\text{Re}\mathbb{E}(X_j(0)X_k(0))$

in the CQF problem of Sect. 6. Also, Lemma 2 can be extended to the more general case $R \succeq 0$ with nonstrict inequalities $\geq$ in (8).

*Example 1* Consider a two-mode QHO ($n = 4$) with a CCR matrix $\Theta := \frac{1}{2} I_2 \otimes J$, where $\otimes$ is the Kronecker product, and

$$J := \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

(42)

spans the space $A_2$. This corresponds to the system variables consisting of two pairs of conjugate position $q_k$ and momentum $-i \partial q_k$ operators on the Schwartz space [44], $k = 1, 2$ (with an appropriately normalized Planck constant [37]). Suppose the QHO has the energy matrix $R := \begin{bmatrix} 3.4048 & 0.0721 & -2.2402 & -2.8090 \\ 3.0478 & 0.8484 & -2.2402 & -2.8090 \\ -2.4614 & 4.7504 & 0.8484 & 4.7504 \\ -2.4614 & 0.8484 & 4.7504 & 4.7504 \end{bmatrix} \succ 0$, so that the frequencies are $\pm 4.3074$, $\pm 0.6540$, and the margin (40) is $\tau_* = 0.7645$. The initial covariance condition (24) is given by

$$\Sigma := \begin{bmatrix} 5.9068 & -2.2359 & -0.8477 & 2.0721 \\ -2.2359 & 4.7534 & 4.6272 & -2.8090 \\ -0.8477 & 4.6272 & 6.7367 & -4.1352 \\ 2.0721 & -2.8090 & -4.1352 & 4.8525 \end{bmatrix},$$

and satisfies the uncertainty relation $\Sigma + i \Theta \succ 0$. Lemma 1 is used in order to compute the discounted second-order moments of the system variables. The dependence of their real parts on the ETH $\tau$ is depicted in Fig. 1.
These graphs show that the interval \( 0 < \tau < 5 \tau_\ast = 3.8225 \) is sufficiently large for the moments to manifest convergence to the infinite-horizon averages: \( \text{Re}E_\infty(XX^T) = \begin{bmatrix} 8.3140 & -4.8573 & 0.3322 & 1.8803 \\ -4.8573 & 5.7935 & 1.5480 & -1.6743 \\ 0.3322 & 1.5480 & 9.3853 & -0.7758 \\ 1.8803 & -1.6743 & -0.7758 & 2.8441 \end{bmatrix} \), as calculated using (41).

## 4 Directly coupled quantum plant and coherent quantum observer

Consider directly coupled quantum plant and coherent quantum observer forming a closed QHO whose Hamiltonian \( H \) is given by

\[
H := \frac{1}{2} \mathcal{X}^T R \mathcal{X}, \quad \mathcal{X} := \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}, \quad \xi := \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_\nu \end{bmatrix},
\]

where \( R \in \mathbb{S}_{n+\nu} \) is the plant–observer energy matrix. Here, \( X_1, \ldots, X_n \) and \( \xi_1, \ldots, \xi_\nu \) are the plant and observer variables, respectively, with both \( n \) and \( \nu \) being even. These quantum variables are time-varying self-adjoint operators on the tensor-product space \( \mathcal{H} := \mathcal{H}_1 \otimes \mathcal{H}_2 \), where \( \mathcal{H}_1, \mathcal{H}_2 \) are initial complex separable Hilbert spaces of the plant and the observer (which can be copies of a common Hilbert space). They are assumed to have a block-diagonal CCR matrix \( \Theta \):

\[
[\mathcal{X}, \mathcal{X}^T] = 2i \Theta, \quad \Theta := \text{diag}(\Theta_k),
\]

where \( \Theta_1 \in \mathbb{A}_n \) and \( \Theta_2 \in \mathbb{A}_\nu \) are nonsingular CCR matrices of the plant and the observer, respectively. For what follows, the plant–observer energy matrix \( R \) in (43) is partitioned as

\[
R := \begin{bmatrix} K & L \\ L^T & M \end{bmatrix}.
\]

Here, \( K \in \mathbb{S}_n \) and \( M \in \mathbb{S}_\nu \) are the energy matrices of the plant and the observer which specify their free Hamiltonians \( H_1 := \frac{1}{2}X^T K X \) and \( H_2 := \frac{1}{2} \xi^T M \xi \). Also, \( L \in \mathbb{R}^{n \times \nu} \) is the plant–observer coupling matrix which parameterizes the interaction Hamiltonian \( H_{12} := \frac{1}{2}(X^T L \xi + \xi^T L^T X) = \text{Re}(X^T L \xi) \), where \( \text{Re}(\cdot) \) applies to operators (and matrices of operators) so that \( \text{Re}N := \frac{1}{2}(N + N^\dagger) \) consists of self-adjoint operators. Hence, the total Hamiltonian in (43) is \( H = H_1 + H_2 + H_{12} \). In view of (43)–(45), the Heisenberg dynamics of the composite system are governed by a linear ODE

\[
\dot{\mathcal{X}} = i[H, \mathcal{X}] = \mathcal{A} \mathcal{X}.
\]
Due to the partitioning of $X$ in (43), the matrix $A \in \mathbb{R}^{(n+\nu) \times (n+\nu)}$ is split into appropriately dimensioned blocks:

\[
A := 2\Theta R = 2 \begin{bmatrix} \Theta_1 K & \Theta_1 L \\ \Theta_2 L^T & \Theta_2 M \end{bmatrix} = \begin{bmatrix} A & BL \\ BL^T & A \end{bmatrix},
\]

so that the ODE (46) can be represented as a set of two ODEs

\[
\dot{X} = AX + B\eta, \quad \dot{\xi} = \alpha \xi + \beta Y,
\]

where

\[
A := 2\Theta_1 K, \quad B := 2\Theta_1, \quad \alpha := 2\Theta_2 M, \quad \beta := 2\Theta_2,
\]

\[
Y := L^T X, \quad \eta := L\xi.
\]

The vector $\eta$ drives the plant variables in (48) and resembles the classical actuator signal. The observer variables in (48) are driven by the plant variables through the vector $Y$ which corresponds to the classical observation output from the plant. However, the quantum mechanical nature of $Y$ and $\eta$ (which consist of time-varying self-adjoint operators on $\mathcal{H}$) makes them qualitatively different from the classical signals [23]. In view of the relation $[Y, L^T] = L^T [X, X^T] L = 2i L^T \Theta_1 L$, following from (44), (50), the outputs $Y_1, \ldots, Y_\nu$ do not commute with each other, in general, and are inaccessible to simultaneous measurement. Since the plant and observer being considered form a fully quantum system which does not involve measurements, $Y$ is not an observation signal in the usual control theoretic sense. In order to emphasize this distinction from the classical case, such observers are referred to as coherent (that is, measurement-free) quantum observers [21, 24, 28, 30, 48, 56]. In addition to the noncommutativity of the dynamic variables, specified by the CCRs (44), the quantum mechanical nature of the setting manifests itself in the fact that the “observation” and “actuation” channels in (50) depend on the same matrix $L$. This coupling between the ODEs (48) is closely related to the Hamiltonian structure $A \in \Theta S_{n+\nu}$ of (47). Therefore, the “quantum information flow” from the plant through $Y$ has a “back-action” effect on the plant dynamics through $\eta$. However, unlike the conventional meaning of this term in the context of quantum measurements, the back-action considered here is caused by the direct coupling of the observer which modifies the plant dynamics.

Assuming that the plant energy matrix $K$ is fixed, the matrices $L, M$ can be varied so as to achieve desired properties for the plant–observer QHO under constraints on the plant–observer coupling. To this end, for a given ETH $\tau > 0$, the observer is called $\tau$-admissible if $A$ in (47) satisfies

\[
\tau < \frac{1}{2 \max(0, \ln r(e^{A\tau}))},
\]

cf. (25) of Lemma 1. The corresponding pairs $(L, M)$ form an open subset of $\mathbb{R}^{n \times \nu} \times \mathbb{S}_\nu$ which depends on $\tau$. In application to the plant–observer system, the discussions of Sect. 3 show that if $R > 0$ (and hence, $A$ has a purely imaginary spectrum), then such
an observer is $\tau$-admissible for any $\tau > 0$. The condition $R > 0$ for (45) is equivalent to

$$K > 0, \quad M > 0, \quad \|A\|_\infty < 1, \quad \Lambda := K^{-1/2}LM^{-1/2},$$

(52)

where the third inequality describes the contraction property of $\Lambda$ whose largest singular value $\|A\|_\infty$ quantifies the “smallness” of the coupling matrix $L$ in comparison with the energy matrices $K$, $M$. If the observer satisfies (52), then any system operator (with appropriate finite moments) in the plant–observer QHO lends itself to the discounted averaging of Sect. 3 for any ETH $\tau > 0$. Also note that the rescaling $\hat{X} := \sqrt{K}X$, $\hat{\xi} := \sqrt{M}\xi$

(53)

of the plant and observer variables leads to a QHO with the CCR matrices $\hat{\Theta}_1 := \sqrt{K}\Theta_1\sqrt{K}$ and $\hat{\Theta}_2 := \sqrt{M}\Theta_2\sqrt{M}$ and the energy matrix $\hat{R} := \begin{bmatrix} I_n & A \\ \Lambda^T & I_\nu \end{bmatrix}$, where $\Lambda$ in (52) plays the role of the coupling matrix.

For what follows, it is assumed that the initial plant and observer variables have a block diagonal matrix of second moments:

$$\Sigma := \text{Re}E(X_0 X_0^T) = \text{diag}(\Sigma_k),$$

(54)

where $\Sigma_k + i\Theta_k \succ 0$. In the zero-mean case $E\mathcal{X}_0 = 0$, this corresponds to $X_0$ and $\xi_0$ being uncorrelated. A physical rationale for the absence of initial correlation is that the observer is prepared independently of the plant and then brought into interaction with the latter at $t = 0$. If the plant and the observer remained uncoupled ($L = 0$), then, in view of Lemma 1 and (54), their variables would remain uncorrelated ($E(X\xi^T) = 0$) and the matrices $P_1 := \text{Re}E_\tau(XX^T)$, $P_2 := \text{Re}E_\tau(\xi\xi^T)$ would be unique solutions of independent ALEs:

$$P_1 = \frac{1}{\tau}L(A_\tau, \Sigma_1), \quad A_\tau := A - \frac{1}{2\tau}I_n,$$

(55)

$$P_2 = \frac{1}{\tau}L(\alpha_\tau, \Sigma_2), \quad \alpha_\tau := \alpha - \frac{1}{2\tau}I_\nu,$$

(56)

where (22) is used, and $A_\tau, \alpha_\tau$ are assumed to be Hurwitz. For arbitrary plant–observer coupling ($L \neq 0$), the matrix

$$\mathcal{P} := \begin{bmatrix} \mathcal{P}_{11} & \mathcal{P}_{12} \\ \mathcal{P}_{21} & \mathcal{P}_{22} \end{bmatrix} := \text{Re}E_\tau(\mathcal{X}\mathcal{X}^T),$$

(57)

which is split into blocks similarly to $\mathcal{A}$ in (47), coincides with the controllability Gramian of the pair $(\mathcal{A}_\tau, \sqrt{\tau^{-1}\Sigma})$ and satisfies the ALE

$$\mathcal{P} = \frac{1}{\tau}L(\mathcal{A}_\tau, \Sigma),$$

(58)
provided the observer is \( \tau \)-admissible in the sense of (51). Here, \( \Sigma \) is the initial covariance condition from (54), and

\[
\mathcal{A}_\tau := \mathcal{A} - \frac{1}{2\tau} I_n + v = \begin{bmatrix} A_\tau & BL \\ \beta L_T & \alpha_\tau \end{bmatrix}
\]  

(59)

is Hurwitz. In the case \( L = 0 \) (when the plant and observer are uncoupled), \( \mathcal{P} \) reduces to the block diagonal matrix

\[
\mathcal{P}_* := \text{diag} (P_k), \quad k = 1, 2
\]

(60)

which is formed from the matrices \( P_1, P_2 \) in (55), (56).

5 Observer back-action on covariance dynamics of the plant

The back-action of the observer can be quantified by the deviation of the covariance dynamics of the plant from those which the plant would have if it were uncoupled from the observer. In view of (57), (60), we will describe this deviation in terms of bilateral bounds for \( \mathcal{P}_{11} - P_1 \) and, more generally, \( \mathcal{P} - \mathcal{P}_* \). To this end, we will use the following technical lemma whose proof is given here for completeness.

Lemma 3  Suppose a matrix \( N \in \left[ \begin{array}{cc} N_{11} & N_{12} \\ N_{21} & N_{22} \end{array} \right] \in \mathbb{S}^+_{2n} \) is split into blocks \( N_{jk} \in \mathbb{R}^{n \times n} \). Then

\[
\pm (N_{12} + N_{21}) \preccurlyeq w N_{11} + \frac{1}{w} N_{22}
\]

(61)

for any \( w > 0 \). Also, if \( N_{11} \succ 0 \) (in addition to \( N 
 \succeq 0 \), then

\[
\pm (N_{12} + N_{21}) \preccurlyeq 2 \sqrt{ \text{tr} (N_{11}^{-1} N_{22}) } N_{11}.
\]

(62)

Proof  Since \( N \succeq 0 \), then

\[
0 \preccurlyeq \left[ \begin{array}{cc} \sqrt{w} I_n & \pm \frac{1}{\sqrt{w}} I_n \\ \pm \frac{1}{\sqrt{w}} I_n & \sqrt{w} I_n \end{array} \right] N \left[ \begin{array}{cc} \sqrt{w} I_n & \pm \frac{1}{\sqrt{w}} I_n \\ \pm \frac{1}{\sqrt{w}} I_n & \sqrt{w} I_n \end{array} \right] = w N_{11} + \frac{1}{w} N_{22} \pm (N_{12} + N_{21}) \text{ for any } w > 0,
\]

which proves (61) (similar inequalities are used, for example, in the proof of [39, Lemma 3]). The parameter \( w \) can be varied so as to “tighten up” the bound (61). More precisely, if \( N_{11} \succ 0 \), then

\[
w N_{11} + \frac{1}{w} N_{22} = \sqrt{N_{11}} (w I_n + \frac{1}{w} N_{11}^{-1/2} N_{22} N_{11}^{-1/2}) \sqrt{N_{11}}
\]

\[
\preccurlyeq \left( w + \frac{1}{w} \text{tr} (N_{11}^{-1} N_{22}) \right) N_{11}.
\]

(63)

Since \( \min_{w>0} \left( w + \frac{1}{w} \text{tr} (N_{11}^{-1} N_{22}) \right) = 2 \sqrt{ \text{tr} (N_{11}^{-1} N_{22}) } \) (achieved at \( w = \sqrt{ \text{tr} (N_{11}^{-1} N_{22}) } \)), a combination of (61) with (63) leads to (62).  \( \square \)
The following lemma will be used to give a precise meaning to the property that the observer output with relatively small mean square values has an appropriately weak effect on the covariance dynamics of the plant.

**Lemma 4** Suppose the directly coupled observer is \( \theta \)-admissible, where

\[
\varsigma := \frac{w \tau}{w + \tau} < \tau < \theta := \frac{m \tau}{m - \tau}
\]  

(64)

are related to the ETH \( \tau \) through auxiliary parameters \( w > 0 \) and \( m > \tau \). Then, the matrix \( \mathcal{P}_{11} \) from (57) satisfies

\[
-L(\mathcal{A}_\varsigma, \frac{1}{w} P_1 + w B \text{Re} \mathbf{E}_\tau(\eta \eta^T) B^T) \preceq \mathcal{P}_{11} - P_1
\]

\[
\preceq L(\mathcal{A}_\theta, \frac{1}{m} P_1 + m B \text{Re} \mathbf{E}_\tau(\eta \eta^T) B^T),
\]

(65)

where \( P_1 \) is given by (55).

**Proof** Due to the inequalities in (64), the \( \theta \)-admissibility of the observer ensures that \( \mathcal{A}_\varsigma, \mathcal{A}_\tau, \mathcal{A}_\theta \) are Hurwitz. From (48), it follows that \( (XX^T)^* = AXX^T + XX^T A^T + B \eta X^T + X \eta^T B^T \). Application of the discounted averaging operator \( \mathbf{E}_\tau \) to the latter ODE and the integration by parts lead to

\[
\frac{1}{\tau} (\mathcal{P}_{11} - \Sigma_1) = \mathcal{A}_\tau \mathcal{P}_{11} + \mathcal{P}_{11} \mathcal{A}_\tau^T + \mathcal{Y},
\]

and hence,

\[
\mathcal{A}_\tau \mathcal{P}_{11} + \mathcal{P}_{11} \mathcal{A}_\tau^T + \frac{1}{\tau} \Sigma_1 + \mathcal{Y} = \mathcal{A}_\tau (\mathcal{P}_{11} - P_1) + (\mathcal{P}_{11} - P_1) \mathcal{A}_\tau^T + \mathcal{Y} = 0.
\]

(66)

Here, the term

\[
\mathcal{Y} := \text{Re} \mathbf{E}_\tau(B \eta X^T + X \eta^T B^T)
\]

(67)

originates from the plant–observer coupling and plays the role of a perturbation to the ALE \( \mathcal{A}_\tau P_1 + P_1 \mathcal{A}_\tau^T + \frac{1}{\tau} \Sigma_1 = 0 \) in (55). By applying (61) of Lemma 3 to the matrix \( N := \text{Re} \mathbf{E}_\tau(\zeta \zeta^T) = \begin{bmatrix} \mathcal{P}_{11} & \text{Re} \mathbf{E}_\tau(\eta \eta^T) B^T \\ B \text{Re} \mathbf{E}_\tau(\eta \eta^T) B^T & \text{Re} \mathbf{E}_\tau(\eta \eta^T) B^T \end{bmatrix} \succeq 0 \) of the real parts of the second-order moments of an auxiliary vector \( \zeta := \begin{bmatrix} X \\ B \eta \end{bmatrix} \), it follows that the matrix \( \mathcal{Y} \) in (67) satisfies

\[
\mathcal{Y} \preceq \frac{1}{w} \mathcal{P}_{11} + w B \text{Re} \mathbf{E}_\tau(\eta \eta^T) B^T \succeq -\mathcal{Y}
\]

(68)

for any \( w > 0 \). Substitution of the second inequality from (68) into (66) leads to

\[
0 \succeq \mathcal{A}_\tau (\mathcal{P}_{11} - P_1) + (\mathcal{P}_{11} - P_1) \mathcal{A}_\tau^T - \frac{1}{w} \mathcal{P}_{11} - w B \text{Re} \mathbf{E}_\tau(\eta \eta^T) B^T
\]

\[
= \mathcal{A}_\varsigma (\mathcal{P}_{11} - P_1) + (\mathcal{P}_{11} - P_1) \mathcal{A}_\varsigma^T - \frac{1}{w} P_1 - w B \text{Re} \mathbf{E}_\tau(\eta \eta^T) B^T,
\]

(69)
where we also use the relation \( \frac{1}{\tau} + \frac{1}{w} = \frac{1}{\zeta} \), which follows from the definition of \( \zeta \) in (64) and implies that \( A_\tau + \frac{1}{2w} I_n = A_\zeta \) in view of (55). The Lyapunov inequality (69) leads to the lower bound for \( \Theta_{11} - P_1 \) in (65). The upper bound in (65) is established in a similar fashion by combining the ALE (66) with the first inequality from (68), except that the parameter \( m := w \) has to satisfy \( m > \tau \) in order to ensure that \( \theta > 0 \) in (64), thus making \( A_\tau + \frac{1}{2m} I_n = A_0 \) Hurwitz.

The parameters \( w, m \) in Lemma 4 can be varied in order to tighten up the bounds (65), similarly to the proof of (62) of Lemma 3. Indeed, let the matrix \( B \text{Re} \mathbf{E}_\tau (\eta\eta^T)B^T \) be small compared to \( P_1 \) in terms of the dimensionless quantity

\[
\kappa := \tau \sqrt{\text{r}(P_1^{-1} B \text{Re} \mathbf{E}_\tau (\eta\eta^T)B^T)},
\]

provided \( P_1 > 0 \). The latter condition is fulfilled, for example, if \( \Sigma_1 > 0 \) in (54). Then, by letting \( w = m = \frac{\tau}{\zeta} \) in (64), \( \zeta = \frac{\tau}{1+\tau} \) and \( \theta = \frac{\tau}{1-\tau} \) become close to \( \tau \) for small values of \( \kappa \), and the bounds (65) behave asymptotically as \( \pm(\Theta_{11} - P_1) \lesssim 2\kappa L A_\tau, P_1) \lesssim 2\kappa r(P_1 \Sigma_1^{-1})P_1 \). The last inequality is obtained by applying (23) to (55). Therefore, Lemma 4 guarantees that \( \Theta_{11} - P_1 \) is small compared to \( P_1 \) (that is, \( r(\Theta_{11} P_1^{-1} - I_n) \ll 1 \)) and the back-action effect is negligible, if the second moments of the observer output \( \eta \) are small in the sense that (70) satisfies \( \kappa \max(1, r(P_1 \Sigma_1^{-1})) \ll 1 \).

Note that Lemma 4 does not employ the relation (50) between the observer output \( \eta \) and the observer variables \( \xi \). We will therefore provide a more accurate bound for the deviation \( \Theta_{11} - P_1 \), which takes into account the whole plant–observer dynamics (48)–(50), including the fact that \( \xi \) is driven by the plant output \( Y \). The formulation of the following theorem employs auxiliary matrices

\[
D_1 := \text{diag}(A_\tau \oplus A_\tau, \alpha_\tau \oplus \alpha_\tau), \quad D_2 := \text{diag}(A_\tau \oplus \alpha_\tau, \alpha_\tau \oplus A_\tau),
\]

\[
E_1 := \begin{bmatrix} I_n \otimes (BL) & (BL) \otimes I_n \\ (BL^T) \otimes I_v & I_v \otimes (BL^T) \end{bmatrix}, \quad E_2 := \begin{bmatrix} I_n \otimes (\beta L^{T}) & (BL) \otimes I_v \\ (BL^T) \otimes I_n & I_v \otimes (BL) \end{bmatrix},
\]

where \( N \oplus Q := N \otimes I + I \otimes Q \) is the Kronecker sum of matrices. The matrices \( D_1, D_2 \) are associated with the \( L \)-independent diagonal blocks of the matrix \( \mathcal{A}_\tau \) from (59), while \( E_1, E_2 \) depend linearly on the coupling matrix \( L \) and are associated with the \( L \)-dependent off-diagonal part of \( \mathcal{A}_\tau \).

**Theorem 2.** Suppose the observer is \( \tau \)-admissible, and \( A_\tau, \alpha_\tau \) in (55), (56) are also Hurwitz. Furthermore, let the plant–observer system have the block-diagonal initial covariance condition (54), and suppose the matrices

\[
\Delta_1 := D_1^{-1} E_1, \quad \Delta_2 := D_2^{-1} E_2,
\]

defined in terms of (71), (72), satisfy

\[
\epsilon := \|[\Delta_1]_{\infty}\|[\Delta_2]_{\infty} < 1.
\]

 Springer
Then, the Frobenius norm \( \| \cdot \|_2 \) of the deviation of \( \mathcal{P} \) in (57) from its value \( \mathcal{P}_\ast \) for uncoupled plant and observer in (60) admits upper bounds

\[
\| \mathcal{P}_{11} - P_1 \|_2 \leq \frac{\epsilon}{1 - \epsilon} \| \mathcal{P}_\ast \|_2, \\
\| \mathcal{P} - \mathcal{P}_\ast \|_2 \leq \sqrt{1 + \| \Delta_1 \|_2^2} \| \Delta_2 \|_\infty \| \mathcal{P}_\ast \|_2.
\]  

(75)  
(76)

**Proof** Although the matrix \( \mathcal{P} \) is symmetric, we will use its full (rather than half-) vectorization \( \text{vec}(\mathcal{P}) \in \mathbb{R}^{(n+\nu)^2} \) [25]. For brevity, the vectorization of a matrix is written as \( \vec{\cdot} \) throughout the proof. The vector \( \vec{\mathcal{P}} \) is obtained by permutating the entries of \( \begin{bmatrix} \vec{P}_{11} \\ \vec{P}_{12} \\ \vec{P}_{22} \end{bmatrix} \) which satisfies the following vectorized form of the ALE (58) in view of (59):

\[
\begin{bmatrix}
A_\tau \oplus A_\tau \\
I_n \otimes (BL) \\
(\beta L^T) \otimes I_n
\end{bmatrix}
\begin{bmatrix}
\vec{P}_{11} \\
\vec{P}_{21} \\
\vec{P}_{22}
\end{bmatrix}
= -\frac{1}{\tau}
\begin{bmatrix}
\vec{\Sigma}_1 \\
\vec{\Sigma}_2
\end{bmatrix}.
\]

The sparsity of its right-hand side results from the block-diagonal structure of (54) and splits the set of linear equations into the non-homogeneous and homogeneous parts

\[
D_1 \begin{bmatrix} \vec{P}_{11} \\ \vec{P}_{22} \end{bmatrix} + E_1 \begin{bmatrix} \vec{P}_{21} \\ \vec{P}_{12} \end{bmatrix} = -\frac{1}{\tau} \begin{bmatrix} \vec{\Sigma}_1 \\ \vec{\Sigma}_2 \end{bmatrix}, \\
D_2 \begin{bmatrix} \vec{P}_{21} \\ \vec{P}_{12} \end{bmatrix} + E_2 \begin{bmatrix} \vec{P}_{11} \\ \vec{P}_{22} \end{bmatrix} = 0,
\]  

(77)

where (71), (72) are used. Since \( A_\tau, \alpha_\tau \) are Hurwitz, \( D_1, D_2 \) are nonsingular. The second equation in (77) for \( \begin{bmatrix} \vec{P}_{21} \\ \vec{P}_{12} \end{bmatrix} \) and substitution of its solution into the first one lead to

\[
\begin{bmatrix} \vec{P}_{21} \\ \vec{P}_{12} \end{bmatrix} = -\Delta_2 \begin{bmatrix} \vec{P}_{11} \\ \vec{P}_{22} \end{bmatrix},
\]

\[
\begin{bmatrix} \vec{P}_{11} \\ \vec{P}_{22} \end{bmatrix} = -\frac{1}{\tau} (D_1 - E_1 \Delta_2)^{-1} \begin{bmatrix} \vec{\Sigma}_1 \\ \vec{\Sigma}_2 \end{bmatrix} = -\frac{1}{\tau} \begin{bmatrix} I_{n^2+\nu^2} - \Delta_1 \Delta_2 \end{bmatrix}^{-1} D_1^{-1} \begin{bmatrix} \vec{\Sigma}_1 \\ \vec{\Sigma}_2 \end{bmatrix}
\]

\[
= (I_{n^2+\nu^2} - \Delta_1 \Delta_2)^{-1} \begin{bmatrix} \vec{P}_1 \\ \vec{P}_2 \end{bmatrix}.
\]  

(78)  
(79)

Here, (73), (74) are used together with the vectorizations \( \vec{P}_1 = -\frac{1}{\tau} (A_\tau \oplus A_\tau)^{-1} \vec{\Sigma}_1 \), \( \vec{P}_2 = -\frac{1}{\tau} (\alpha_\tau \oplus \alpha_\tau)^{-1} \vec{\Sigma}_2 \) of the solutions of the ALEs (55), (56). Since \( \Delta := \Delta_1 \Delta_2 \) is a contraction, with \( \| \Delta \|_\infty \leq \epsilon \), the perturbation expansion of the matrix inverse [15]...
in (79) yields
\[
\begin{bmatrix}
\tilde{P}_{11} - \tilde{P}_1 \\
\tilde{P}_{22} - \tilde{P}_2
\end{bmatrix}
= \Delta(I_{n^2 + n^2} - \Delta)^{-1}
\begin{bmatrix}
\tilde{P}_1 \\
\tilde{P}_2
\end{bmatrix}
= \sum_{k=1}^{+\infty} \Delta^k
\begin{bmatrix}
\tilde{P}_1 \\
\tilde{P}_2
\end{bmatrix}
\leq \sum_{k=1}^{+\infty} \epsilon^k
\begin{bmatrix}
\tilde{P}_1 \\
\tilde{P}_2
\end{bmatrix}
= \epsilon \frac{1}{1 - \epsilon} |\tilde{\mathcal{P}}_*|.
\] (80)

Since the vectorization preserves the Frobenius norm, then $|\tilde{\mathcal{P}}_*| = \sqrt{\sum_{k=1}^{2} \|P_k\|^2_2} = \|\mathcal{P}_*\|_2$ in view of (60), and hence, (80) implies (75). To prove (76), note that the triangle inequality and (80) lead to
\[
\begin{bmatrix}
\tilde{P}_{11} \\
\tilde{P}_{22}
\end{bmatrix}
\leq \begin{bmatrix}
\tilde{P}_1 \\
\tilde{P}_2
\end{bmatrix} + \begin{bmatrix}
\tilde{P}_{11} - \tilde{P}_1 \\
\tilde{P}_{22} - \tilde{P}_2
\end{bmatrix}
\leq \frac{\|\mathcal{P}_*\|_2}{1 - \epsilon},
\] (81)
where $\begin{bmatrix}
\tilde{P}_1 \\
\tilde{P}_2
\end{bmatrix} = \|\mathcal{P}_*\|_2$ is used again. A combination of (81) with (78) implies that
\[
\begin{bmatrix}
\tilde{P}_{21} \\
\tilde{P}_{12}
\end{bmatrix}
\leq \Delta_2 \frac{\|\mathcal{P}_*\|_2}{1 - \epsilon}.\] (82)

The orthogonal decomposition $\mathcal{P}_* = \text{diag}_{k=1,2}(\mathcal{P}_{kk} - P_k) + \begin{bmatrix} 0 & P_{12} \\ P_{21} & 0 \end{bmatrix}$ and (80), (82) lead to $\|\mathcal{P}_*\|_2 = \sqrt{\sum_{k=1}^{2} \|P_k\|^2_2} = \|\mathcal{P}_*\|_2$ which establishes (76) in view of (74).

Since Theorem 2 employs the standard (rather than weighted) Frobenius norm $\|\cdot\|_2$, it would be physically more meaningful to apply the theorem to covariance dynamics of the rescaled plant and observer variables (53). Alternatively, Theorem 2 can be reformulated in terms of a weighted version of the norm. In this case, $\|\mathcal{P}_*\|_2$ is replaced with $\|S(\mathcal{P}_* - \mathcal{P}_*)S\|_2$, where $S := \text{diag}(\sqrt{K}, \sqrt{M})$. The standard Frobenius norm is used in (75), (76) for simplicity of formulation. There is a parallel between the proof of Theorem 2 and the arguments underlying the small-gain theorem (see, for example, [9] and references therein). Similar bounds for the observer back-action are obtained in the frequency domain as outlined below. From (48), (50), it follows that the Laplace transforms $\tilde{X}, \tilde{\xi}$ of the plant and observer vectors $X, \xi$, given by (32), are related by
\[
\tilde{X}(s) = F(s)(X_0 + BL\tilde{\xi}(s)), \quad \tilde{\xi}(s) = \Phi(s)(\xi_0 + \beta L^T\tilde{X}(s)),
\] (83)
see Fig. 2.

Here, $F, \Phi$ are the plant and observer transfer functions, which are given by
\[
F(s) := (sI_n - A)^{-1}, \quad \Phi(s) := (sI_n - \alpha)^{-1}
\] (84)
Fig. 2 A block diagram of (83), with the initial values $X_0$ and $\xi_0$ shown as fictitious external inputs. A small-gain-theorem argument applies when the coupling matrix $L$ is relatively small in accordance with (31) and do not depend on the coupling matrix $L$. From (83), (84), it follows that the Laplace transform of the combined vector $\mathcal{X}$ of the plant and observer variables in (43) is related to its initial value $\mathcal{X}_0$ by

$$\tilde{\mathcal{X}}(s) = \begin{bmatrix} \tilde{X}(s) \\ \tilde{\xi}(s) \end{bmatrix} = \begin{bmatrix} G(s) \\ 0 \end{bmatrix} \mathcal{X}_0$$

through the transfer function

$$G(s) := \begin{bmatrix} I_n & -F(s)BL \\ -\Phi(s)\beta L^T & I_v \end{bmatrix}^{-1} \begin{bmatrix} F(s) & 0 \\ 0 & \Phi(s) \end{bmatrix}. \quad (85)$$

By applying (29) to the plant–observer system, the matrix $P$ in (57) is represented as:

$$P = \frac{1}{2\pi \tau} \text{Im} \int_{\text{Res} = \frac{1}{2\tau}} G(s)(\Sigma + i\Theta)G(s)^* ds. \quad (86)$$

The function $G$ in (85) differs from $\text{diag}(F, \Phi)$ by the factor

$$\begin{bmatrix} I_n & -T_1(s) \\ -T_2(s) & I_v \end{bmatrix}^{-1} \begin{bmatrix} (I_n - T_1(s)T_2(s))^{-1} & (I_n - T_1(s)T_2(s))^{-1}T_1(s) \\ (I_v - T_2(s)T_1(s))^{-1}T_2(s) & (I_v - T_2(s)T_1(s))^{-1} \end{bmatrix}$$

which is close to $I_{n+v}$ for the relevant values of $s \in \mathbb{C}$ (with $\text{Res} = \frac{1}{2\tau}$), provided $L$ is small in the sense of the quantities

$$\gamma_k := \sup_{\omega \in \mathbb{R}} \left\| T_k \left( \frac{1}{2\tau} + i\omega \right) \right\|_\infty, \quad k = 1, 2. \quad (87)$$

The latter are “discounted” versions of the $H_\infty$ Hardy space norm for the transfer functions $T_1(s) := F(s)BL$ and $T_2(s) := \Phi(s)\beta L^T$ which depend linearly on $L$. Therefore, the frequency-domain representation (86) can be used together with the parameters (87) in order to obtain bounds for the deviation of $P$ from the matrix $P_\ast = \frac{1}{2\pi \tau} \text{Im} \int_{\text{Res} = \frac{1}{2\tau}} \text{diag}(F(s)\Gamma_1 F(s)^*, \Phi(s)\Gamma_2\Phi(s)^*) ds$ under the small-gain condition $\gamma_1 \gamma_2 < 1$ for the loop in Fig. 2. Here, $\Gamma_k := \Sigma_k + i\Theta_k$ are the initial second moment matrices for the plant and observer variables, which, in accordance with (54), form the matrix $E(\mathcal{X}_0, \mathcal{X}_0^T) = \text{diag}_{k=1,2}(\Gamma_k)$. 

\(\text{Springer}\)
6 Discounted mean square optimal coherent quantum filtering problem

Let the plant energy matrix be fixed and satisfy $K \succ 0$, which is sufficient for the set of $\tau$-admissible observers to be nonempty for any given $\tau > 0$. In particular, this set contains observers with $L = 0$ and arbitrary $M \succ 0$ (in which case $R = \text{diag}(K, M) \succ 0$), or more generally, observers satisfying (52). Consider a CQF problem

$$Z := \mathbb{E}_\tau Z \longrightarrow \min$$

of minimizing a quadratic cost over the plant–observer coupling matrix $L$ and the observer energy matrix $M$ subject to the constraint (51). Here, $\tau > 0$ is a given ETH for the discounted averaging (15) applied to the quantum process

$$Z := E^T E + \lambda \eta^T \Pi \eta = \mathcal{X}^T \mathcal{C}^T \mathcal{C} \mathcal{X}.$$  \hspace{1cm} (89)

The latter is a time-varying self-adjoint operator on the plant–observer space $\mathcal{H}$ which is defined in terms of the vectors $\mathcal{X}, \eta$ from (43), (50), and

$$E := S_X - S_2 \xi = S \mathcal{X}, \quad S := \begin{bmatrix} S_1 & -S_2 \end{bmatrix}, \quad \mathcal{C} := \begin{bmatrix} S_1 & -S_2 \
0 & \sqrt{\lambda \Pi} L \end{bmatrix}. \hspace{1cm} (90)$$

Here, $S_1 \in \mathbb{R}^{p \times n}$, $S_2 \in \mathbb{R}^{p \times \nu}$, $\Pi \in \mathbb{S}_n$ are given matrices, with $\Pi \succ 0$, which, together with a given scalar parameter $\lambda > 0$, determine the matrix $\mathcal{C} \in \mathbb{R}^{(p+n) \times (n+\nu)}$ (with the first block-row $S \in \mathbb{R}^{p \times (n+\nu)}$) and its dependence on the coupling matrix $L$. The matrix $S_1$ specifies linear combinations of the plant variables of interest which are to be approximated by given linear functions of the observer variables specified by the matrix $S_2$. Accordingly, the vector $E$ in (90) (consisting of $p$ time-varying self-adjoint operators on $\mathcal{H}$) is interpreted as an estimation error. In addition to the discounted mean square $\mathbb{E}_\tau (E^T E)$ of this error, the cost $\mathcal{Z}$ in (88) involves a quadratic penalty $\mathbb{E}_\tau (\eta^T \Pi \eta)$ for the observer back-action on the covariance dynamics of the plant (see Lemma 4), with $\lambda$ being the relative weight of this penalty in $\mathcal{Z}$. In fact, $\mathcal{Z}$ is organized as the Lagrange function for a related CQF problem of minimizing the discounted mean square of the estimation error subject to an additional weighted mean square constraint on the plant–observer coupling:

$$\mathbb{E}_\tau (E^T E) \longrightarrow \min, \quad \mathbb{E}_\tau (\eta^T \Pi \eta) \leq r.$$

(91)

In this formulation, $\lambda$ plays the role of a Lagrange multiplier which is found so as to make the solution of (88) saturate the constraint in (91) for a given threshold $r > 0$. In a particular case of $S_2 = 0$, the CQF problem (88)–(90) is a quantum mechanical analogue of the LQR problem [23] in view of the analogy between the observer output $\eta$ and classical actuation signals discussed in Sect. 4. The presence of the quantum expectation of a nonlinear function of system variables in (88) and the optimization requirement make this setting different from the time-averaged approach of [34, 36].
Substitution of (89) in (88) allows the cost functional to be expressed in terms of the matrix $P$ from (57) as:

$$\mathcal{L} = \langle C^T C, E_\tau (P X^T) \rangle = \langle C^T C, P \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the Frobenius inner product of matrices. Under the assumptions of Theorem 2, a combination of (92) with the Cauchy–Bunyakovsky–Schwarz inequality and the bound (76) leads to

$$|E_\tau (E^T E) - \langle S^T S, P_* \rangle| = |\langle S^T S, P - P_* \rangle| \leq \|S^T S\|_2 \|P - P_*\|_2 \leq \|S^T S\|_2 \sqrt{\frac{1 + \|\Delta_1\|_\infty^2}{1 - \epsilon}} \|\Delta_2\|_\infty \|P_*\|_2,$$

relating the discounted mean square of the estimation error with the plant–observer coupling strength quantified by $\|\Delta_1\|_\infty$, $\|\Delta_2\|_\infty$, $\epsilon$ from (73), (74). Here, $S$ is the first block-row of the matrix $C$ in (90), so that $\|S^T S\|_2 = \|SS^T\|_2 = \sqrt{\text{Tr}(\sum_{k=1}^2 S_k S_k^T)^2}$, and $\langle S^T S, P_* \rangle = \sum_{k=1}^2 \text{Tr}(S_k P_k S_k^T)$ in view of the block-diagonal structure of the matrix $P_*$ in (60). Therefore, (93) implies

$$E_\tau (E^T E) \geq \sum_{k=1}^2 \text{Tr}(S_k P_k S_k^T) - \sqrt{\text{Tr}\left(\left(\sum_{k=1}^2 S_k S_k^T\right)^2\right) \frac{1 + \|\Delta_1\|_\infty^2}{1 - \epsilon}} \|\Delta_2\|_\infty \|P_*\|_2,$$

which provides a lower bound for the mean square of the estimation error (and becomes an equality if $L = 0$). This bound depends on the coupling matrix $L$ only through $\|\Delta_1\|_\infty$, $\|\Delta_2\|_\infty$, $\epsilon$ and shows that $L$ has to be sufficiently large in order to make $E_\tau (E^T E)$ smaller than $\sum_{k=1}^2 \text{Tr}(S_k P_k S_k^T)$ by a given amount. At the same time, the plant–observer coupling should be weak enough to avoid severe back-action of the observer on the plant. Therefore, the parameter $\lambda$ in the CQF problem (88)–(90) quantifies a compromise between these conflicting requirements (of minimizing the estimation error and reducing the back-action).

### 7 First-order necessary conditions of optimality

Using the variational approach of [47, 48], the following theorem provides the first-order necessary conditions of optimality for the CQF problem (88)–(90). Their formulation employs the Hankelian

$$\mathcal{E} := \begin{bmatrix} \mathcal{E}_{11} & \mathcal{E}_{12} \\ \mathcal{E}_{21} & \mathcal{E}_{22} \end{bmatrix} = \mathcal{Q} \mathcal{P},$$

where $\mathcal{Q}$ and $\mathcal{P}$ are matrices related to the CQF problem.
associated with $\mathcal{P}$ in (57) and the observability Gramian $\mathcal{Q}$ of $(\mathcal{S}_\tau, \mathcal{C})$ which is a unique solution of the ALE

$$\mathcal{Q} := \begin{bmatrix} \mathcal{Q}_{11} & \mathcal{Q}_{12} \\ \mathcal{Q}_{21} & \mathcal{Q}_{22} \end{bmatrix} = L(\mathcal{S}^T_{\tau}, \mathcal{C}^T \mathcal{C}).$$  \hspace{1cm} (95)

The matrices $\mathcal{E}$, $\mathcal{Q}$ are split into appropriately dimensioned blocks $(\cdot)_{jk}$ similar to the matrix $\mathcal{P}$ in (57), with $(\cdot)_{j\bullet}$ the $j$th block-row and $(\cdot)_{\bullet k}$ the $k$th block-column of the matrices. These matrices enter the Frechet derivatives $\partial_L \mathcal{Z}, \partial_M \mathcal{Z}$ of the cost (88) (as a function $(L, M) \mapsto \mathcal{Z}$ of the plant–observer coupling matrix $L \in \mathbb{R}^{n \times \nu}$ and the observer energy matrix $M \in \mathcal{S}_{\nu}$), which are computed below and equated to zero in order to obtain the conditions of stationarity of $\mathcal{Z}$ over $L, M$ as the first-order necessary conditions of optimality.

**Theorem 3** Suppose the plant energy matrix satisfies $K \succ 0$, and the directly coupled observer is $\tau$-admissible in the sense of (51). Then, the observer is a stationary point of the CQF problem (88)–(90) if and only if the Hankelian $\mathcal{E}$ in (94) and the controllability Gramian $\mathcal{P}$ in (57) satisfy

$$\Theta_1 \mathcal{E}_{12} - \mathcal{E}_{21} \Theta_2 = \frac{\lambda}{2} \mathcal{P} L \mathcal{P}_{22}, \quad \Theta_2 \mathcal{E}_{22} - \mathcal{E}_{22} \Theta_2 = 0. \hspace{1cm} (96)$$

**Proof** By using (58) and the duality $L(\mathcal{S}_\tau, \cdot)^T = L(\mathcal{S}^T_{\tau}, \cdot)$, it follows that the cost $\mathcal{Z}$ in (92) is representable in terms of the observability Gramian $\mathcal{Q}$ from (95) as:

$$\mathcal{Z} = \frac{1}{\tau} \mathcal{E}^T \mathcal{C}, \mathcal{L}(\mathcal{S}_\tau, \Sigma) = \frac{1}{\tau} \mathcal{L}(\mathcal{S}^T_{\tau}, \mathcal{C}^T \mathcal{C}), \Sigma = \frac{1}{\tau} (\mathcal{Z}, \Sigma). \hspace{1cm} (97)$$

Here, the adjoint $(\cdot)^T$ of linear operators on matrices is in the sense of the Frobenius inner product. With the matrix $\mathcal{S}_\tau$ in (59) being Hurwitz due to the $\tau$-admissibility constraint (51), the representation (97) shows that $\mathcal{Z}$ inherits a smooth dependence on $L, M$ from $\mathcal{Q}$. The latter is a composite function $(L, M) \mapsto (\mathcal{S}, \mathcal{C}) \mapsto \mathcal{Q}$ whose first variation is

$$\delta \mathcal{Q} = L(\mathcal{S}^T_{\tau}, (\delta \mathcal{S})^T \mathcal{Q} + \mathcal{Q} (\delta \mathcal{C})^T \mathcal{C} + \mathcal{C}^T \delta \mathcal{C}), \hspace{1cm} (98)$$

where the ALE (95) is used, and the first variations of the matrices $\mathcal{S}$ in (47) and $\mathcal{C}$ in (90) with respect to $L, M$ are

$$\delta \mathcal{S} = 2 \Theta \begin{bmatrix} 0 & \delta L \\ \delta L^T & \delta M \end{bmatrix}, \delta \mathcal{C} = \begin{bmatrix} 0 & 0 \\ 0 & \sqrt{\lambda \mathcal{P} \delta L} \end{bmatrix}. \hspace{1cm} (99)$$

By combining the duality argument above with (98), (99), it follows that the first variation of $\mathcal{Z}$ in (97) can be computed as (see, for example, [47] for similar calculations)
\[
\delta \mathcal{L} = \frac{1}{\tau} \left( \mathbf{L}(\alpha, (\delta \alpha)^T \mathcal{L} + \mathcal{D} \delta \alpha + (\delta \mathcal{E})^T \mathcal{E} + \mathcal{E}^T \delta \mathcal{E}), \Sigma \right)
\]

\[
= \left( (\delta \alpha)^T \mathcal{L} + \mathcal{D} \delta \alpha + (\delta \mathcal{E})^T \mathcal{E} + \mathcal{E}^T \delta \mathcal{E}, \mathcal{P} \right)
\]

\[
= 2 \langle \mathcal{E}, \delta \alpha \rangle + 2 \langle \mathcal{E}^T \delta \mathcal{E}, \mathcal{P} \rangle
\]

\[
= -4 \left( \Theta \mathcal{E}, \begin{bmatrix} 0 & \delta L \\ \delta L^T & \delta M \end{bmatrix} \right) + 2 \langle \mathcal{E} \mathcal{P}, \begin{bmatrix} 0 & \sqrt{\lambda \Pi} \delta L \\ 0 \end{bmatrix} \rangle
\]

\[
= -8 \langle \mathcal{S}(\Theta \mathcal{E})_{12}, \delta L \rangle - 4 \langle \mathcal{S}(\Theta \mathcal{E})_{22}, \delta M \rangle + 2(\sqrt{\lambda \Pi} L \mathcal{P}_{22}, \sqrt{\lambda \Pi} \delta L)
\]

\[
= 2 \langle \lambda \Pi L \mathcal{P}_{22} - 4 \mathcal{S}(\Theta \mathcal{E})_{12}, \delta L \rangle - 4 \langle \mathcal{S}(\Theta \mathcal{E})_{22}, \delta M \rangle.
\] (100)

Here, \( S(N) := \frac{1}{2} (N + N^T) \) is the matrix symmetrizer, so that

\[
\mathcal{S}(\Theta \mathcal{E}) = \frac{1}{2} (\Theta \mathcal{E} - \mathcal{E}^T \Theta) = \frac{1}{2} \begin{bmatrix} \Theta_1 \mathcal{E}_{11} - \mathcal{E}_{11}^T \Theta_1 & \Theta_1 \mathcal{E}_{12} - \mathcal{E}_{21}^T \Theta_2 \\ \Theta_2 \mathcal{E}_{21} - \mathcal{E}_{12}^T \Theta_1 & \Theta_2 \mathcal{E}_{22} - \mathcal{E}_{22}^T \Theta_2 \end{bmatrix}.
\] (101)

A combination of (100) with (101) leads to the partial Frechet derivatives of \( \mathcal{L} \) on the Hilbert spaces \( \mathbb{R}^{n \times v} \), \( \mathcal{S}_v : \partial_L \mathcal{L} = 2(\lambda \Pi L \mathcal{P}_{22} - 4 \mathcal{S}(\Theta \mathcal{E})_{12}) = 2(\lambda \Pi L \mathcal{P}_{22} - 2(\Theta_1 \mathcal{E}_{12} - \mathcal{E}_{21}^T \Theta_2)) \) and \( \partial_{\mathcal{E}} \mathcal{L} = -4 \mathcal{S}(\mathcal{E}_{22}) = -2(\Theta_2 \mathcal{E}_{22} - \mathcal{E}_{22}^T \Theta_2) \). These derivatives vanish (that is, \( \mathcal{L} \) is stationary with respect to \( L, M \)) if and only if (96) are satisfied.

In view of (101), the first-order optimality conditions (96) for the observer are equivalent to the existence of a matrix \( N \in \mathbb{S}_n \) such that

\[
\Theta \mathcal{E} - \mathcal{E}^T \Theta = \frac{1}{2} \begin{bmatrix} N & \lambda \Pi L \mathcal{P}_{22} \\ \lambda \Pi L \mathcal{P}_{22}^T & 0 \end{bmatrix}.
\] (102)

Here, the zero block corresponds to the second condition in (96), whereby \( \mathcal{E}_{22} \) is skew-Hamiltonian in the sense of the symplectic structure specified by \( \Theta_2^{-1} : \mathcal{E}_{22} \in \Theta_2^{-1} \mathbb{A}_v \).

A quantum probabilistic interpretation of the conditions (96) is that, for any such observer, the process \( \vartheta \), given by

\[
\vartheta := \begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} := \Theta \mathcal{X}, \quad \vartheta_j := \Theta_j \mathcal{X}, \quad j = 1, 2,
\] (103)

and consisting of \( n + v \) self-adjoint operators (which are special linear combinations of the plant and observer variables), satisfies the covariance relations

\[
\mathbf{E}_\tau (\vartheta_1 \xi^T + \mathcal{X} \vartheta_2^T) = \Theta \mathcal{X} \mathbf{E}_\tau (\mathcal{X} \xi^T) - \mathbf{E}_\tau (\mathcal{X} \mathcal{X}^T) \mathcal{E}_2 \Theta_2
\]

\[
= \Theta \mathcal{X} \left[ \mathcal{P}_{22} + i \Theta_2 \right] - (\mathcal{P} + i \Theta) \mathcal{E}_2 \Theta_2 = \Theta \mathcal{E}_2 - \mathcal{E}_2 \Theta_2
\]
\[
\begin{bmatrix}
\frac{1}{2} \Pi L P_{22} \\
0
\end{bmatrix} = \begin{bmatrix}
\frac{1}{2} \Pi \text{Re} \mathbf{E}_r (\eta \xi^T) \\
0
\end{bmatrix}.
\] (104)

Here, we have used the identities \( \varepsilon_{jk} = \mathcal{D}_j \mathcal{P}_k \), \( \varepsilon_{jk}^T = \mathcal{P}_k \mathcal{D}_j \), which follow from (94) and the symmetry of the Gramians \( \mathcal{P} \), \( \mathcal{D} \) in (57), (95). In particular, (104) implies that \( \vartheta_2 \) in (103) and \( \xi \) are uncorrelated in the sense that \( \mathbb{E}_T (\vartheta_2 \xi^T + \xi \vartheta_2^T) = 0 \). This is a quantum counterpart of the corresponding property for the estimation error and the state estimate in the classical Kalman filter [23]. If \( P_{22} > 0 \), then, in view of the assumption \( \Pi > 0 \), (96) implies that the optimal coupling matrix is
\[
L = \frac{2}{\lambda} \Pi^{-1} (\Theta_1 \varepsilon_{12} - \varepsilon_{21}^T \Theta_2) P^{-1}_{22}.
\] (105)

The relations (58), (95), (105) have to be complemented with an equation for the optimal observer matrix \( M \), which is less straightforward and is considered in the next section.

### 8 Lie-algebraic representation of optimality conditions

The Gramians \( \mathcal{P} \), \( \mathcal{D} \) from (57), (95) give rise to
\[
P := \mathcal{P} \Theta^{-1}, \quad Q := \Theta \mathcal{D}
\] (106)

belonging to the same subspace \( \Theta S_{n+\nu} \) of Hamiltonian matrices as \( \mathcal{A} \) in (47). Here, \( P \in \Theta S_{n+\nu} \) since \( \Theta^{-1} \mathcal{P} \Theta^{-1} \in S_{n+\nu} \). The above ALEs and optimality conditions are represented by the following lemma in terms of the linear space \( \Theta S_{n+\nu} \), with the commutator \([\cdot, \cdot]\), which is a Lie algebra [10, 43]. Its formulation employs the Hamiltonian matrix
\[
D := [Q, P] = \Theta \mathcal{D} \mathcal{P} \Theta^{-1} - \mathcal{P} \mathcal{D} = (\Theta \varepsilon - \varepsilon^T \Theta) \Theta^{-1},
\] (107)

which (for any \( \tau \)-admissible observer) is related to (102) due to (94), (106) and the symmetry of the Gramians \( \mathcal{P} \), \( \mathcal{D} \).

**Lemma 5** The ALEs (58), (95) and the optimality conditions (96) for the CQF problem (88)–(90) admit a Lie-algebraic form through the Hamiltonian matrices (106):
\[
[\mathcal{A}, P] = \frac{1}{\tau} (P - \Sigma \Theta^{-1}), \quad [\mathcal{A}, Q] = \Theta \varepsilon^T \varepsilon - \frac{1}{\tau} Q,
\] (108)
\[
D_{12} = \frac{\lambda}{2} \Pi L P_{22}, \quad D_{22} = 0,
\] (109)

where \( D_{12}, D_{22} \) are the blocks of \( D \in \Theta S_{n+\nu} \) in (107).

**Proof** The Hamiltonian structure of \( \mathcal{A} \) in (47) implies that \( \mathcal{A}^T = -\Theta^{-1} \mathcal{A} \Theta \), and hence,

\( \square \) Springer
\[ A_\tau P + P A_\tau^T = A P - \Theta^{-1} A \Theta - \frac{1}{\tau} P \]
\[ = (A P - \Theta^{-1} A \Theta - \frac{1}{\tau} P) \Theta, \quad (110) \]
\[ A_\tau^T Q + Q A_\tau = A_\tau^T Q - \frac{1}{\tau} Q \]
\[ = -\Theta^{-1} A \Theta Q + Q \Theta - \frac{1}{\tau} Q = -\Theta^{-1}(A \Theta + \frac{1}{\tau} Q), \quad (111) \]

where (59), (106) are used. Substitution of (110), (111) into the ALEs (58), (95) leads to their Lie-algebraic representations (108). By substituting (102) into (107), considering the second block-column
\[ D_2 = \frac{1}{\tau} \begin{bmatrix} \lambda \Pi L P_{22} & \Theta^{-1} \\ 0 & \Theta \end{bmatrix} \]
and using \( P_{22}, \Theta^{-1} \) follows that (96) take the form (109).

The solutions of (108) admit the representation
\[ P = (I - \tau \text{ad}_{\Phi})^{-1}(\Sigma \Theta^{-1}), \quad Q = \tau(I + \tau \text{ad}_{\Phi})^{-1}(\Theta \mathcal{C}^T \mathcal{C}), \quad (112) \]
with \( I \) the identity operator on \( \Theta S_{n+\nu} \). The resolvents \( (I \pm \tau \text{ad}_{\Phi})^{-1} \) are well defined since the \( \tau \)-admissibility (51) implies that the spectrum of \( \text{ad}_{\Phi} \) on \( \Theta S_{n+\nu} \) is in the strip \( \{ z \in \mathbb{C} : |\text{Re}z| < \frac{1}{\tau} \} \).

**Lemma 6** The optimal coupling matrix \( L \) in (105) is expressed in terms of \( P, Q \) from (106) as:
\[ L = \frac{2}{\lambda} \Pi^{-1} D_{12} P_{22}^{-1}, \quad (113) \]

provided \( P_{22} \succ 0 \), where \( D \) is given by (107). The optimal energy matrix \( M \) of the observer satisfies
\[ \frac{1}{2} \left( \frac{1}{\tau} [\Sigma \Theta^{-1}, Q]_{12} + [\Theta \mathcal{C}^T \mathcal{C}, P]_{12} \right) + D_{11} \Theta_{1} L - \Theta_{1} K D_{12} + D_{12} \Theta_{2} M = 0. \quad (114) \]

**Proof** If \( P_{22} \succ 0 \), (113) follows directly from the first optimality condition in (109). In order to establish (114), note that (107), (108) involve pairwise commutators of the Hamiltonian matrices \( A, P, Q \in \Theta S_{n+\nu} \). Hence, the Jacobi identity [43] and the antisymmetry of the commutator lead to
\[ 0 = [[P, A], Q] + [[A, Q], P] + [[Q, P], A] \]
\[ = \frac{1}{\tau} [\Sigma \Theta^{-1} - P, Q] + \left[ \Theta \mathcal{C}^T \mathcal{C} - \frac{1}{\tau} Q, P \right] + [D, A] \]
\[ = \frac{1}{\tau} [\Sigma \Theta^{-1}, Q] + [\Theta \mathcal{C}^T \mathcal{C}, P] + [D, A] \quad (115) \]
for any \( \tau \)-admissible observer, where neither of the optimality conditions (109) has been used. By substituting \( A \) from (47) into (115) and considering the \((\cdot)_{12}\) block of the resulting Hamiltonian matrix, it follows that

\[
\frac{1}{\tau} [\Sigma \Theta^{-1}, Q]_{12} + [\Theta E^T \Theta, P]_{12} + 2(D_{11} \Theta_1 L + D_{12} \Theta_2 M - \Theta_1 (KD_{12} + LD_{22})) = 0. \tag{116}
\]

Now, the second optimality condition in (109) makes the corresponding term in (116) vanish, thus leading to (114).

The proof of Lemma 6 shows that (114) holds for any \( \tau \)-admissible stationary point of the CQF problem regardless of \( \mathcal{P}_{22} > 0 \). Furthermore, (114) is a linear equation for the optimal observer energy matrix \( M \). This allows \( M \) to be expressed in terms of \( P, Q \) in (106) in the case of equal plant and observer dimensions \( n = v \). In this case, the observer will be called nondegenerate if \( P, D \) in (106), (107) satisfy

\[
\mathcal{P}_{22} > 0, \quad \det D_{12} \neq 0. \tag{117}
\]

The above results lead to the following necessary conditions of optimality for such observers.

**Theorem 4** Suppose the plant and observer dimensions are equal: \( n = v \). Then for any nondegenerate observer, which is a stationary point of the CQF problem (88)–(90) under the assumptions of Theorem 3, the coupling and energy matrices are related by (113) and by

\[
M = \Theta^{-1}_2 D_{12}^{-1} \left( \Theta_1 KD_{12} - D_{11} \Theta_1 L - \frac{1}{2} \left[ \frac{1}{\tau} [\Sigma \Theta^{-1}, Q]_{12} + [\Theta E^T \Theta, P]_{12} \right] \right) \tag{118}
\]

to the matrices \( P, Q \) from (106) satisfying the ALEs (108).

**Proof** The first of the conditions (117) makes the representation (113) applicable, with \( \det L \neq 0 \) in view of the second condition in (117). The latter allows (114) to be solved uniquely for the observer energy matrix \( M \) as in (118).

The first line of (118) is organized as a similarity transformation which would relate the Hamiltonian matrices \( \Theta_1 K, \Theta_2 M \) if there were no additional terms on the right-hand side of (118). In that case, the transformation matrix \( D_{12} \) would preserve the Hamiltonian structure if it were symplectic in the generalized sense that \( D_{12} \Theta_2 D_{12}^T = \Theta_1 \). In combination with the ALEs (58), (95) (or their Lie-algebraic form (108)), the relations (113), (118) of Lemma 6 and Theorem 4 provide a set of algebraic equations for finding the matrices \( L, M \) of a nondegenerate observer among stationary points of the CQF problem (88)–(90).
9 Observers with autonomous estimation error dynamics

In view of the complicated structure of the equations of Sects. 7 and 8 for an optimal observer, consider the CQF problem for a special class of observers with autonomous dynamics of the estimation error \( E \) in (90). More precisely, suppose the observer is such that

\[
S\mathcal{A} = \hat{\mathcal{A}} S
\]

for some matrix \( \hat{\mathcal{A}} \in \mathbb{R}^{p \times p} \), where \( S \) is given by (90). In combination with (46), the relation (119) leads to the ODE

\[
\dot{E} = S\hat{\mathcal{A}}\dot{X} = \hat{\mathcal{A}} S\dot{X} = \hat{\mathcal{A}} E.
\]

These autonomous dynamics preserve the CCRs for the estimation error:

\[
[E, E^T] = 2i\hat{\Theta}, \quad \hat{\Theta} := S\Theta S^T = \sum_{k=1}^{2} S_k\Theta_k S_k^T.
\]

Indeed, from (119), (121) and the Hamiltonian property \( \mathcal{A} \in \Theta S_{n+\nu} \), it follows that

\[
\hat{\mathcal{A}}\hat{\Theta} + \hat{\Theta}\hat{\mathcal{A}}^T = \hat{\mathcal{A}} S\Theta S^T + S\Theta S^T \hat{\mathcal{A}}^T
\]

\[
= S(\mathcal{A}\Theta + \Theta\mathcal{A}^T)S^T = 0.
\]

Therefore, if the CCR matrix \( \hat{\Theta} \in \mathcal{A}_p \) in (121) is nonsingular, (122) implies that \( \hat{\mathcal{A}} \in \hat{\Theta}S_p \). Now, let the plant and observer have equal dimensions \( n = \nu \) and identical CCR matrices

\[
\Theta_0 := \Theta_1 = \Theta_2,
\]

with \( \Theta_0 \in \mathcal{A}_n \), \( \det \Theta_0 \neq 0 \). Also, suppose the estimation error \( E \) in (90) has the same dimension \( p = n \) and is specified by

\[
S_0 := S_1 = S_2 \in \mathbb{R}^{n \times n}, \quad \det S_0 \neq 0.
\]

Then, the process \( E \) reduces to

\[
E = S_0(X - \xi),
\]

and its CCR matrix in (121) is nonsingular:

\[
\hat{\Theta} = 2S_0\Theta_0 S_0^T.
\]

Since \( E^T E = (X - \xi)^T S_0^T S_0(X - \xi) \) in view of (125), the matrix \( S_0^T S_0 > 0 \) specifies the relative importance of the plant variables in the CQF problem (88), (89).
Lemma 7  Under the conditions (123), (124), the estimation error (125) acquires the autonomous dynamics (120) due to (119) for some \( \mathcal{A} \in \mathbb{R}^{n \times n} \) if and only if the observer has the same energy matrix as the plant and a symmetric coupling matrix:

\[
M = K, \quad L = L^T.
\]  
(127)

For any such observer, the matrix \( \mathcal{A} \) is found uniquely as

\[
\mathcal{A} = 2\hat{\Theta} \hat{R},
\]  
(128)

where \( \hat{\Theta} \) is the estimation error CCR matrix in (126), and

\[
\hat{R} := \frac{1}{2} S_0^{-T} (K - L) S_0^{-1}
\]  
(129)

is a real symmetric matrix of order \( n \).

**Proof**  A combination of (47), (123), (124) leads to

\[
S \mathcal{A} = 2S_0 \Theta_0 \left[ K - L^T \quad L - M \right], \quad \mathcal{A} S = \mathcal{A} S_0 \left[ I_n \quad -I_n \right].
\]  
(130)

Therefore, since \( \det S_0 \neq 0 \), the fulfillment of (119) for some \( \mathcal{A} \in \mathbb{R}^{n \times n} \) is equivalent to \( K - L^T = M - L \), that is,

\[
M - K = L - L^T.
\]  
(131)

Since the left-hand side of (131) is a symmetric matrix, while the right-hand side is antisymmetric, and only the zero matrix has these two properties simultaneously (that is, \( \mathbb{S}_n \cap A_n = \{0\} \)), then (131) holds if and only if \( L, M \) satisfy (127). In this case, (119), (126), (130) yield \( \mathcal{A} = 2S_0 \Theta_0 (K - L) S_0^{-1} = \hat{\Theta} S_0^{-T} (K - L) S_0^{-1} \), leading to (128), with \( \hat{R} \) given by (129). \( \square \)

The observer in Lemma 7 replicates the quantum plant, except that it has a different initial space and, in general, different initial covariance conditions in (54). The structure (127) of such observers does not depend on particular matrices \( \Theta_0, S_0 \). In view of (120), (128), the entries of the estimation error \( E \) in (125) evolve in time as system variables of a QHO with the CCR matrix \( \hat{\Theta} \) in (126) and the energy matrix \( \hat{R} \) in (129). Without additional constraints on the coupling matrix \( L \) (apart from its symmetry in (127)), \( \hat{R} \) can be ascribed any given value in \( \mathbb{S}_n \) by an appropriate choice of \( L \). However, large values of \( L \) are penalized by the second term of the cost functional in (89). A solution of the CQF problem (88) in this class of observers is as follows.

**Theorem 5**  In the framework of Lemma 7 under the conditions (123), (124) and \( \mathcal{P}_{22} \succ 0 \), an optimal coupling matrix \( L \in \mathbb{S}_n \) for the observer with autonomous estimation error dynamics satisfies

\[
L = \frac{8}{\lambda} \Pi^{-1} \left( - \mathcal{P}_{22} \Pi^{-1}, S(\Theta \mathcal{E}(\Theta))^{-1} \right) \Pi^{-1}.
\]  
(132)
Proof In view of (127), the observer energy matrix \( M = K \) remains fixed, and, due to the symmetry of \( L \), the first variation (100) of the cost functional in the proof of Theorem 3 reduces to

\[
\delta Z = 2 \langle S(\lambda \Pi \mathcal{P}_{22} - 4S(\Theta E)_{12}), \delta L \rangle.
\]

Hence,

\[
\partial L Z = 2 S(\lambda \Pi \mathcal{P}_{22} - 4S(\Theta E)_{12}) = \lambda (\mathcal{P}_{22} \Pi + \Pi \mathcal{P}_{22}) - 8S(\Theta E)_{12}) = \lambda (\mathcal{P}_{22} \Pi^{-1} \tilde{L} + \tilde{L} \Pi^{-1} \mathcal{P}_{22}) - 8S(\Theta E)_{12}),
\]

where

\[
\tilde{L} := \Pi L \Pi
\]

inherits its symmetry from \( L \) and \( \Pi \). From (133), it follows that \( \partial L Z = 0 \) is equivalent to \( \tilde{L} \) being a unique solution of the ALE

\[
\tilde{L} = \frac{8}{\lambda} (\mathcal{P}_{22} \Pi^{-1}, S(\Theta E)_{12})),
\]

where \( -\mathcal{P}_{22} \Pi^{-1} \) is isospectral to \( -\Pi^{-1/2} \mathcal{P}_{22} \Pi^{-1/2} \prec 0 \) and hence, is Hurwitz. A combination of (134) with (135) leads to (132). \( \square \)

The right-hand side of (132) is a nonlinear composite function of the coupling matrix \( L \) and a scalar parameter

\[
\mu := \frac{1}{\lambda} > 0
\]

(136)

which is assumed to be sufficiently small) and has the form

\[
L_\mu = \mu f(\mu, L_\mu).
\]

The computation of the function \( f \) involves the solution of the ALEs (58), (95) for the Gramians \( \mathcal{P}, \mathcal{Q} \) with the matrix

\[
\mathcal{A} = 2 \begin{bmatrix} \Theta_0 K & \Theta_0 L \\ \Theta_0 L & \Theta_0 K \end{bmatrix},
\]

followed by computing the Hankelian \( E \) in (94) and solving the ALE (135). The parameter \( \mu \) in (136) enters \( f \) only through

\[
E^T E = S^T S + \begin{bmatrix} 0 & 0 \\ 0 & \frac{1}{\mu} L \Pi L \end{bmatrix}
\]

in the ALE (95). The smallness of \( \mu \) corresponds to large values of \( \lambda \) (high penalization of the observer back-action on the plant). For all sufficiently small \( \mu > 0 \) and \( L \in S_n \), the function \( f \) is Frechet differentiable, and this smoothness is inherited by \( L_\mu \) in
The differentiation of (137) with respect to \( \mu \) (as fictitious time) leads to the ODE

\[
\partial_\mu L_\mu = (\mathcal{J} - \mu \partial_L f)^{-1} (f + \mu \partial_\mu f),
\]

with the initial condition \( L_0 = 0 \), where \( \mathcal{J} \) is the identity operator on \( \mathbb{S}_n \), and \( \partial_L f \) is the appropriate partial Frechet derivative of \( f \). The initial-value problem (140) describes a homotopy method for numerical solution of the CQF problem, similar to [26] (see also, [45]). The right-hand side of (140) is well defined for all \((\mu, L)\) in a small neighbourhood of \((0, 0)\). Its computation can be implemented by using the vectorization of the Frechet derivatives of solutions of ALEs [41, 47] in application to the ALEs (58), (95) (or their Lie-algebraic forms (112)) and (135). The details of these calculations are tedious and omitted for brevity. The weak-coupling (or high-penalization) asymptotic behaviour of the matrix \( L_\mu \) is described below.

**Theorem 6** Suppose the uncoupled observer satisfies \( P_2 > 0 \) in (56). Then, as \( \lambda \to +\infty \) in (89), the optimal coupling matrix in Theorem 5 behaves asymptotically as:

\[
L_\mu \sim \mu L', \quad \text{as} \quad \mu \to 0+,
\]

where \( L' \in \mathbb{S}_n \) is a unique solution of the ALE

\[
L' = 2 \Pi^{-1} L (-P_2 \Pi^{-1}, P_+ \mathcal{D}_0 \Theta_0 - \Theta_0 \mathcal{D}_0 P_+) \Pi^{-1}.
\]

Here, the matrix \( P_+ \) is related to the second-moment matrices (55), (56) for the uncoupled plant and observer variables by

\[
P_+ := P_1 + P_2 = \frac{1}{\tau} L(A_\tau, \Sigma_1 + \Sigma_2),
\]

with

\[
A_\tau := A - \frac{1}{2\tau} I_n, \quad A = 2 \Theta_0 K.
\]

Also, the matrix

\[
\mathcal{D}_0 := S_0^T \hat{\mathcal{D}} S_0
\]

in (142) is associated with a unique solution \( \hat{\mathcal{D}} \) of the ALE

\[
\hat{\mathcal{D}} := L(\mathcal{A}_\tau^T, I_n), \quad \mathcal{A}_\tau := \mathcal{A} - \frac{1}{2\tau} I_n, \quad \mathcal{A} = 2 S_0 \Theta_0 K S_0^{-1}.
\]

\( \square \) Springer
**Proof** From the representation (137) of (132) (or from (140)), it follows that (141) holds with

\[ L' := \left. \frac{\partial}{\partial \mu} L_{\mu} \right|_{\mu=0} = f(0,0), \]  

(147)

where the initial condition \( L_0 = 0 \) is also used. Here,

\[ f(0,0) = 8\mathcal{P}^{-1} L(-P_2 \mathcal{P}^{-1}, S(S(\Theta E)_{12})) \mathcal{P}^{-1} \]  

(148)

is associated with the uncoupled plant and observer. In this case, (138) reduces to \( \mathcal{A} = 2I_2 \otimes (\Theta_0 K) \) with a purely imaginary spectrum due to \( K > 0 \), and \( P_1, P_2 \) in the block-diagonal controllability Gramian (60) satisfy the ALEs (55), (56) with the common matrix \( A_\tau = \alpha_\tau \) in (144), so that their sum is given by (143). In the limit of uncoupled plant and observer, \( \frac{1}{\mu} L_{\mu} \Pi L_{\mu} \sim \mu f(0,0) \Pi f(0,0) \to 0 \) as \( \mu \to 0^+ \), whereby (139) leads to \( \mathcal{C}^T \mathcal{C} = S^T S \) at \( \mu = 0 \), and the ALE (95) for the observability Gramian \( \mathcal{D} \) takes the form:

\[ \mathcal{A}_T \mathcal{D} + \mathcal{D} \mathcal{A}_T + S^T S = 0. \]  

(149)
Fig. 4 The dependence of the back-action penalty term \( \lambda E_T(\eta^T \Pi \eta) \) on the parameter \( \mu \) in Example 2

The property (119) of the observers under consideration implies that

\[
S A T = S A - \frac{1}{2T} S = A S - \frac{1}{2T} S = A T S \quad \text{and hence,} \quad (149) \text{ admits a lower-rank solution}
\]

\[
Q = S^T \hat{Q} S.
\]  

(150)

Indeed, substitution in (149) yields \( S^T (A T \hat{Q} + \hat{Q} A T + I_n) S = 0 \). Therefore, (146) makes (150) a unique solution of the ALE (149), since \( \hat{A} \) is isospectral to \( 2 \Theta_0 K \) with a purely imaginary spectrum (whereby \( \hat{A} \) is Hurwitz). Since \( S = \begin{bmatrix} 1 & -1 \end{bmatrix} \otimes S_0 \), the Hankelian takes the form

\[
\mathcal{E} = S^T \hat{Q} S \text{ diag}_{k=1,2}(P_k) = \begin{bmatrix} \mathcal{D}_0 P_1 - \mathcal{D}_0 P_2 \\ -\mathcal{D}_0 P_1 & \mathcal{D}_0 P_2 \end{bmatrix}, \text{ with } \mathcal{D}_0 \text{ given by (145), and hence, } S(S(\Theta \mathcal{E})_{12}) = \frac{1}{4} (P_+ \mathcal{D}_0 \Theta_0 - \Theta_0 \mathcal{D}_0 P_+). \text{ Substitution in (148), (147) leads to (142).} 
\]

\( \Box \)

Example 2 Let the plant and observer be one-mode QHOs (\( n = \nu = 2 \)), with the CCR matrix \( \Theta_0 := \frac{1}{2} J \) (corresponding to the position-momentum pair, with \( J \) given by (42), and an energy matrix \( K := \begin{bmatrix} 2.7604 & -1.7564 \\ -1.7564 & 2.4982 \end{bmatrix} > 0 \). The frequencies of such a QHO are \( \pm 1.9522 \), and the margin (40) is \( \tau_* = 0.2561 \). The initial covariance conditions in (54) for the plant and observer (prepared independently) are

\[
\Sigma_1 = \begin{bmatrix} 4.1400 & -2.4687 \\ -2.4687 & 4.3641 \end{bmatrix},
\]

\[
\Sigma_2 = \begin{bmatrix} 2.2174 & 1.3387 \\ 1.3387 & 2.4695 \end{bmatrix}
\]

and satisfy the uncertainty relation constraints \( \Sigma_k + i \Theta_0 \not\succ 0 \). With the ETH chosen to be \( \tau = 4.0614 \gg \tau_* \), the second-moment matrices of
The dependence of the entries of the coupling matrix $L_\mu$ on the parameter $\mu$ in Example 2 is shown in Fig. 5.

For the CQF problem (88), the observer back-action penalty matrix $\Pi$ in (89) and the weighting matrix $S_0$ in the estimation error (125) are given by

\[
\Pi = \begin{bmatrix}
1.2907 & 0.9694 \\
0.9694 & 3.7716
\end{bmatrix},
S_0 = \begin{bmatrix}
-1.7389 & 0.2192 \\
0.0170 & 1.0458
\end{bmatrix}.
\]

The mean square of the estimation error for the uncoupled observer is $\text{Tr}(S_0 P_1 + S_0^T P_1) = 46.8634$. The mean square value $E_\tau(E^T E)$ for the optimal observer in the CQF problem (88) (subject to the autonomous estimation error dynamics) is shown in Fig. 3 for a range of values of $\mu$ in (136). This is a strictly decreasing function of $\mu$, whose computation (along with the optimal observers) was carried out using the homotopy method starting from the uncoupled observer (at $\mu = 0$). The observer back-action penalty term $\lambda E_\tau(\eta^T \Pi \eta)$ is shown in Fig. 4. The entries of the corresponding optimal coupling matrix $L_\mu$ are presented in Fig. 5. The calculation of the matrix $L'$, which specifies their asymptotic behaviour as $\mu \to 0+$ according to Theorem 6, yielded $L' = \begin{bmatrix}
-0.7297 & -1.7445 \\
-1.7445 & 1.1737
\end{bmatrix}$. For $0 < \mu \leq 5$ (that is, $\lambda \geq 0.2$), the plant–observer energy matrix $\begin{bmatrix} K & L_\mu \\ L_\mu^T & K \end{bmatrix}$ remained positive definite, so that the system variables retained oscillatory behaviour, justifying the discounted averaging approach.
10 Conclusion

We have considered the computation of discounted averages with exponentially decaying weights for moments of system variables for QHOs, including the mean square functionals, both in the state space and frequency domain. For a quantum plant and a quantum observer in the form of directly coupled QHOs, we have obtained small-gain-theorem bounds for the back-action of the observer on the covariance dynamics of the plant in terms of the plant–observer coupling. We have considered a CQF problem of minimizing the discounted mean square value of the estimation error together with a penalty on the observer back-action. First-order necessary conditions of optimality have been obtained for this problem in the form of algebraic matrix equations involving two coupled ALEs. We have applied Lie-algebraic techniques to these equations and discussed a solution of the CQF problem in the case of autonomous estimation error dynamics, including the homotopy method for its implementation. These results have been illustrated by numerical experiments.

Acknowledgements The authors thank the anonymous reviewer and Associate Editor for useful comments.

Author Contributions Inspired by the earlier results of IRP on directly coupled coherent quantum observers with time averaging [34], IGV developed discounted performance criteria for QHOs, small-gain-theorem bounds for the observer back-action, the mean square optimal CQF setting with penalized back-action, optimality conditions, the Lie-algebraic approach and homotopy algorithm for this problem, carried out numerical experiments and drafted the manuscript. IGV and IRP had extensive discussions and agreed on the final version of the paper.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Arnold VI (1989) Mathematical methods of classical mechanics, 2nd edn. Springer, New York
2. Bernstein DS, Haddad WM (1989) LQG control with an $H^\infty$ performance bound: a Riccati equation approach. IEEE Trans Autom Control 34(3):293–305
3. Besicovitch AS (1954) Almost periodic functions. Dover, New York
4. Billingsley P (1968) Convergence of probability measures. Wiley, New York
5. Blackwell D (1965) Discounted dynamic programming. Ann Math Stat 36(1):226–235
6. Breuer H-P, Petruccione F (2006) The theory of open quantum systems. Clarendon Press, Oxford
7. Caldeira AO, Leggett AJ (1981) Influence of dissipation on quantum tunneling in macroscopic systems. Phys Rev Lett 46(4):211–214
8. Cushen CD, Hudson RL (1971) A quantum-mechanical central limit theorem. J Appl Probab 8(3):454–469
9. D’Helon C, James MR (2006) Stability, gain, and robustness in quantum feedback networks. Phys Rev A 73(5):053803
10. de Gosson M (2006) Symplectic geometry and quantum mechanics. Birkhäuser, Basel
11. Edwards SC, Belavkin VP (2005) Optimal quantum filtering and quantum feedback control. arXiv:quant-ph/0506018v2
12. Feller W (1971) An introduction to probability theory and its applications, vol II, 2nd edn. Wiley, New York
13. Gardiner CW, Zoller P (2004) Quantum noise, 3rd edn. Springer, Berlin
14. Gough J, James MR (2009) Quantum feedback networks: Hamiltonian formulation. Commun Math Phys 287:1109–1132
15. Higham NJ (2008) Functions of matrices. SIAM, Philadelphia
16. Holevo AS (1991) Quantum stochastic calculus. J Math Sci 56(5):2609–2624
17. Holevo AS (2001) Statistical structure of quantum theory. Springer, Berlin
18. Horn RA, Johnson CR (2007) Matrix analysis. Cambridge University Press, New York
19. Hudson RL, Parthasarathy KR (1984) Quantum Ito’s formula and stochastic evolutions. Commun Math Phys 93:301–323
20. Isserlis L (1918) On a formula for the product-moment coefficient of any order of a normal frequency distribution in any number of variables. Biometrika 12:134–139
21. James MR, Nurdin HI, Petersen IR (2008) $H^\infty$ control of linear quantum stochastic systems. IEEE Trans Autom Control 53(8):1787–1803
22. Janson S (1997) Gaussian Hilbert spaces. Cambridge University Press, Cambridge
23. Kwakernaak H, Sivan R (1972) Linear optimal control systems. Wiley, New York
24. Lloyd S (2000) Coherent quantum feedback. Phys Rev A 62(2):022108
25. Magnus JR (1988) Linear structures. Oxford University Press, New York
26. Marinov M, Bertrand P (1985) A homotropy algorithm for solving coupled Riccati equations. Optim Control Appl Methods 6(4):351–357
27. Meyer P-A (1995) Quantum probability for probabilists, 2nd edn. Springer, Berlin
28. Miao Z, James MR (2012) Quantum observer for linear quantum stochastic systems. In: 51st IEEE conference on decision control, Maui, Hawaii, USA, December 10–13, pp 1680–1684
29. Nielsen MA, Chuang IL (2000) Quantum computation and quantum information. Cambridge University Press, Cambridge
30. Nurdin HI, James MR, Petersen IR (2009) Coherent quantum LQG control. Automatica 45:1837–1846
31. Parthasarathy KR (1992) An introduction to quantum stochastic calculus. Birkhäuser, Basel
32. Parthasarathy KR (2010) What is a Gaussian state? Commun Stoch Anal 4(2):143–160
33. Petersen IR (2016) Quantum linear systems theory. Open Autom Control Syst J 8:67–93
34. Petersen IR (2014) A direct coupling coherent quantum observer. In: IEEE MSC 2014, Nice/Antibes, France, 8–10 October 2014, pp 1960–1963
35. Petersen IR (2017) Time averaged consensus in a direct coupled coherent quantum observer network. Control Theory Technol 15(3):163–176
36. Petersen IR, Huntington EH (2015) A possible implementation of a direct coupling coherent quantum observer. In: Australian control conference, Gold Coast, Australia, pp 105–107
37. Sakurai JJ (1994) Modern quantum mechanics. Addison-Wesley, Reading
38. Shaiju AJ, Petersen IR (2012) A frequency domain condition for the physical realizability of linear quantum systems. IEEE Trans Autom Control 57(8):2033–2044
39. Sichani AK, Vladimirov IG, Petersen IR (2014) Robust mean square stability of open quantum stochastic systems with Hamiltonian perturbations in a Weyl quantization form. In: Australian control conference Canberra, 17–18 November 2014, pp 83–88
40. Sichani AK, Vladimirov IG, Petersen IR (2017) A numerical approach to optimal coherent quantum LQG controller design using gradient descent. Automatica 85:314–326
41. Skelton RE, Iwasaki T, Grigoriadis KM (1998) A unified algebraic approach to linear control design. Taylor & Francis, London
42. Sussmann HJ, Willems JC (1997) 300 years of optimal control: from the brachystochrone to the maximum principle. Control Syst 17(3):32–44
43. Varadarajan VS (1984) Lie groups, Lie algebras, and their representations. Springer, New York
44. Vladimirov VS (2002) Methods of the theory of generalized functions. Taylor & Francis, London
45. Vladimirov IG, Petersen IR (2010) Hardy–Schatten norms of systems, output energy cumulants and linear quadro-quartic Gaussian control. In: 19th international symposium on mathematical theory of networks and systems, Budapest, Hungary, July 5–9, pp 2383–2390
46. Vladimirov IG, Petersen IR (2011) A dynamic programming approach to finite-horizon coherent quantum LQG control. In: Australian control conference, Melbourne, pp 357–362
47. Vladimirov IG, Petersen IR (2013) A quasi-separation principle and Newton-like scheme for coherent quantum LQG control. Syst Control Lett 62(7):550–559
48. Vladimirov IG, Petersen IR (2013) Coherent quantum filtering for physically realizable linear quantum plants. In: European control conference, Zurich, Switzerland, 17–19 July, pp 2717–2723
49. Vladimirov IG (2015) A transverse Hamiltonian variational technique for open quantum stochastic systems and its application to coherent quantum control. In: IEEE multi-conference on systems and control, Sydney, Australia, pp 29–34
50. Vladimirov IG (2015) Weyl variations and local sufficiency of linear observers in the mean square optimal coherent quantum filtering problem. In: Australian control conference, Gold Coast, Australia, pp 93–98
51. Vladimirov IG, Petersen IR (2016) Directly coupled observers for quantum harmonic oscillators with discounted mean square cost functionals and penalized back-action. In: 2016 IEEE conference on Norbert Wiener in the 21st century (21CW). University of Melbourne, Australia, pp 78–83
52. Vladimirov IG (2017) A phase-space formulation and Gaussian approximation of the filtering equations for nonlinear quantum stochastic systems. Control Theory Technol 15(3):177–192
53. Walls DF, Milburn GJ (1994) Quantum optics. Springer, Berlin
54. Williamson J (1936) On the algebraic problem concerning the normal forms of linear dynamical systems. Am J Math 58(1):141–163
55. Williamson J (1937) On the normal forms of linear canonical transformations in dynamics. Am J Math 59(3):599–617
56. Wiseman HM, Milburn GJ (1994) All-optical versus electro-optical quantum limited feedback. Phys Rev A 49(5):4110–4125

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.