

State-Space Control Fundamentals (00860721)

lecture no. 9

Leonid Mirkin

The Stephen B. Klein Faculty of Aerospace Engineering
Technion—IIT



1/48

Outline

Observers in the presence disturbances and measurement noise

Linear-quadratic estimation

Coloring linear-quadratic estimation

Linear Quadratic Gaussian (LQG) problem

LQG controller design (optional self-study)

2/48

State reconstruction: more realistic setup

Consider

$$\begin{cases} \dot{x}(t) = Ax(t) + B_d d(t) + Bu(t), & x(0) = x_0 \\ y(t) = Cx(t) + n(t) \end{cases}$$

with the goal to reconstruct the state x from measurements y . Additions to the previously studied state reconstruction setup:

- disturbance d , represents effects of the environment unmeasured
- measurement noise n , represents sensors inaccuracies unmeasured

These are qualitatively different exogenous inputs. Roughly,

- d reflects inaccuracies in our model $\implies$ measurements are important
- n reflects measurements accuracy $\implies$ process model is important

Intuitively, that shall affect the choice of the observer gain L , for

- low-gain observers rely more upon the model think of $\dot{\hat{x}} = A\hat{x} + Bu$
- high-gain observers rely more upon the quality of measurements

3/48

Effect exogenous inputs on error dynamics

Full-order observer

$$\dot{\hat{x}} = A\hat{x} + Bu - L(y - C\hat{x})$$

results now in the error equation for $\epsilon = x - \hat{x}$

$$\dot{\epsilon} = (A + LC)\epsilon + B_d d + Ln$$

Reduced-order observer (to reconstruct $z = C_z x$)

$$\begin{cases} \dot{\hat{\xi}} = (C_z + LC)AT_1\hat{\xi} + (C_z + LC)(A(T_2 - T_1L)y + Bu) \\ \hat{z} = \hat{\xi} - Ly \end{cases}$$

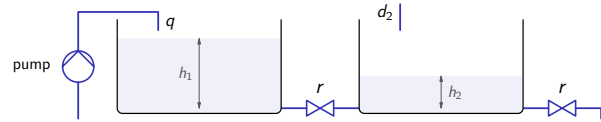
results now in the error equation for $\epsilon_z = z - \hat{z}$

$$\begin{cases} \dot{\epsilon}_{\xi} = (C_z + LC)AT_1\epsilon_{\xi} + C_z B_d d - (C_z + LC)A(T_2 - T_1L)n \\ \epsilon_z = \epsilon_{\xi} + Ln \end{cases}$$

In both cases error dynamics is affected by both d and n .

4/48

Example: double-tank



Linearized dynamics

$$\begin{cases} \dot{x}(t) = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} d_2(t), & x(0) = x_0 \\ y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) + n(t) \end{cases}$$

with

- measured (control) input $u(t) = 0.5 \sin(0.5\pi t) \mathbb{1}(t)$
- disturbance $d_2(t) = 0.05 \mathbb{1}(t)$
- measurement noise $n(t) = 0.01 \sin(4\pi t) \mathbb{1}(t)$

5/48

Example: designs considerations

Full-order observer: pole assignment to roots of $\hat{\chi}_{cl}(s) = s^2 + 1.6\hat{\omega}_n s + \hat{\omega}_n^2$ for $\hat{\omega}_n \in \{1, 8, 64\}$, results in (observer gains increase with $\hat{\omega}_n$)

$$L = - \begin{bmatrix} -1.4 & 9.8 & 99.4 \\ 2.8 & 43.4 & 3896.2 \end{bmatrix}$$

Reduced-order observer: pole assignment to $-1/\tau$ by gain $L = 2 - 1/\tau$; the transfer functions of the systems $T_{\epsilon_z d} : d \mapsto \epsilon_z$ and $T_{\epsilon_z n} : n \mapsto \epsilon_z$ (here the error $\epsilon_z := x_2 - \hat{x}_2$) can be calculated analytically in this case, as

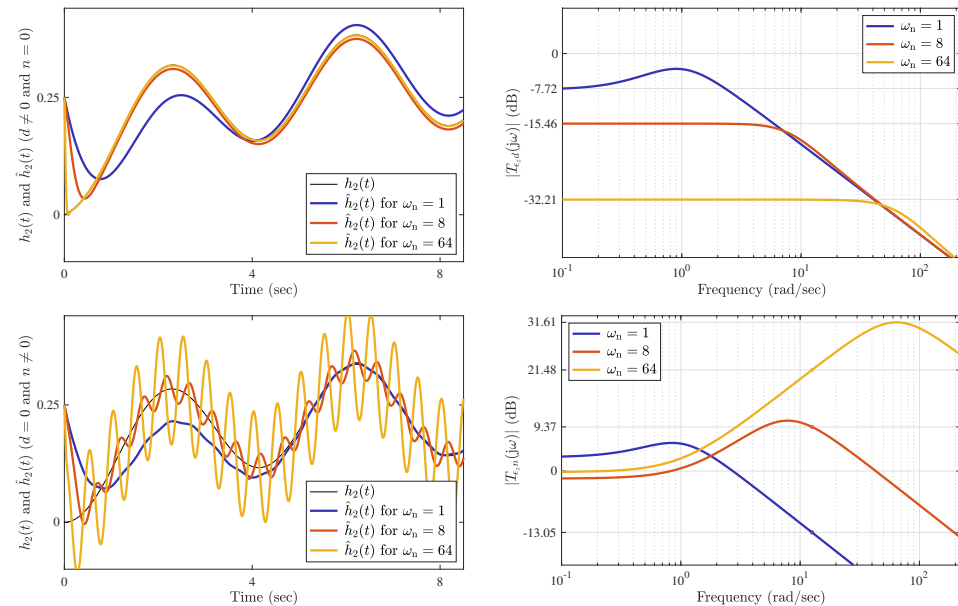
$$T_{\epsilon_z d}(s) = \frac{\tau}{\tau s + 1} \quad \text{and} \quad T_{\epsilon_z n}(s) = \frac{(2\tau - 1)s + \tau - 1}{\tau s + 1} = -G_{\hat{x}_2 y}(s)$$

Simulations are carried out with $\tau \in \{1.5 + \sqrt{1.25}, 0.272313, 0.025\}$, where

1. provides $|T_{\epsilon_z n}(j\omega)| = 1.5 + \sqrt{1.25}$ for all ω
2. provides same $|T_{\epsilon_z n}(j\omega)| = 1.5 + \sqrt{1.25}$ for $\omega = 4\pi$ (noise frequency)
3. yields about the same $|T_{\epsilon_z d}(0)|$ as the full-order observer for $\hat{\omega}_n = 64$

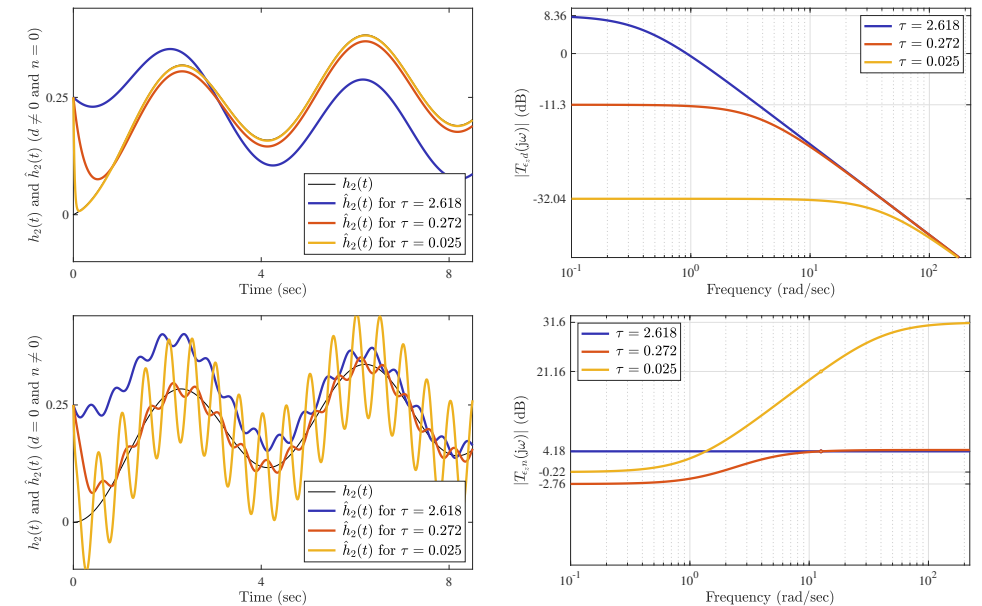
6/48

Example: full-order observer simulations



7/48

Example: reduced-order observer simulations



8/48

Example: discussion

Some conclusions:

- disturbance attenuation and noise sensitivity might be conflicting
- high-gain observers tend to
 - attenuate disturbances better, but
 - at the expense of a higher noise sensitivity
- reduced-order observers tend to be substantially more sensitive to noise especially at high frequencies; compare two **high-gain designs**, which have about the same $|T_{\epsilon_z d}(0)|$ and $\sup_{\omega} |T_{\epsilon_z n}(j\omega)|$, but non-decaying $|T_{\epsilon_z n}(j\omega)|$ at high frequencies

These conclusions are fairly general, although

- pole-placement techniques are not sufficiently transparent in trading off sensitivities to plant disturbances d and measurement noise n .

Way more transparency is offered by optimization-based approaches, where intensities of d and n are explicit tuning parameters.

9/48

Outline

Observers in the presence disturbances and measurement noise

Linear-quadratic estimation

Coloring linear-quadratic estimation

Linear Quadratic Gaussian (LQG) problem

LQG controller design (optional self-study)

10/48

Problem formulation

Let

$$\begin{cases} \dot{x}(t) = Ax(t) + B_w w(t) + B_u u(t), & x(0) = 0 \\ y(t) = Cx(t) + D_w w(t) \end{cases}$$

for given $A \in \mathbb{R}^{n \times n}$, $B_u \in \mathbb{R}^{n \times m_u}$, $B_w \in \mathbb{R}^{n \times m_w}$, $C \in \mathbb{R}^{p \times n}$, & $D_w \in \mathbb{R}^{p \times m_w}$. We need to reconstruct the state x from the measurements y and u so that the reconstructed state $\hat{x}$ is as “close” to x as possible. The linear-quadratic approach is to measure the closeness in terms of the cost function

$$\mathcal{J} = \sum_{i=1}^{m_w} \int_0^\infty \|x(t) - \hat{x}_i(t)\|^2 dt,$$

where $\hat{x}_i := \hat{x}$ under $w = e_i \delta$, where e_i is the i th standard basis of $\mathbb{R}^{m_w}$ and δ is the Dirac delta.

Remark. As we already saw, it is possible to reconstruct x so that the reconstruction error is not affected by the measured input u . So the problem can be formulated with $u = 0$ w.l.o.g.

11/48

Rationale for the choice of w

A scalar signal $x : \mathbb{R} \rightarrow \mathbb{R}$ is said to be **white** if

- its frequency spectrum $|X(j\omega)|^2 = \sigma_x = \text{const}$ for all ω , with $\sigma_x \geq 0$ called the **intensity** of x . Thus,

- white signal has all harmonics equally represented in it cf. white light

It may be safe to say that white signals do not exist in nature. However, it's a convenient *abstraction* used as a cornerstone for generating more realistic signals (via coloring transformations, we shall see that).

Remark. White signals are conventionally introduced and interpreted from the stochastic point of view, as white random processes. Yet such approach goes far beyond the scope of this course (and the math required to define continuous-time white signals rigorously).

Thus,

$$\int_0^\infty \|x(t) - \hat{x}_i(t)\|^2 dt$$

is the reconstruction error energy for a unit-intensity (i.e. normalized) white w acting on the i th input channel. Then $\mathcal{J}$ averages these energies.

12/48

Problem data

In

$$\begin{cases} \dot{x}(t) = Ax(t) + B_w w(t) + B_u u(t), & x(0) = 0 \\ y(t) = Cx(t) + D_w w(t) \end{cases}$$

signals $B_w w$ and $D_w w$ represent plant disturbances and measurement noise, respectively.

Remark. If $D_w B_w' = 0$, then components of w affecting the state (disturbances) and output (noises) equations are mutually independent. The full row rank of D_w , i.e. $D_w D_w' > 0$, would manifest that all measurement channels are corrupt by noise.

Special case: the already familiar

$$\begin{cases} \dot{x}(t) = Ax(t) + B_d d(t) + B_u u(t), & x(0) = 0 \\ y(t) = Cx(t) + n(t) \end{cases}$$

corresponds to $w = \begin{bmatrix} d \\ n \end{bmatrix}$ and $B_u = B$, $B_w = \begin{bmatrix} B_d & 0 \end{bmatrix}$ and $D_w = \begin{bmatrix} 0 & I \end{bmatrix}$.

Special case: the choice $B_w = \begin{bmatrix} B_d \Sigma_d^{1/2} & 0 \end{bmatrix}$ and $D_w = \begin{bmatrix} 0 & \Sigma_n^{1/2} \end{bmatrix}$ reflects weights added to d and n to emphasize relative importance of components.

13/48

Assumptions

A₁: (C, A) is detectable

necessary for a stable observer to exist

A₂: $D_w D_w' > 0$

guarantees that the optimal observer gain is bounded

A₃: $\begin{bmatrix} A - \lambda I & B_w \\ C & D_w \end{bmatrix}$ has full row rank for all $\lambda \in \{z \in \mathbb{C} \mid \operatorname{Re} z = 0\}$.

guarantees that the optimal observer exists and is unique

They guarantee that the CARE

$$\bar{Y} A' + A \bar{Y} + B_w B_w' - (B_w D_w' + \bar{Y} C')(D_w D_w')^{-1} (D_w B_w' + C \bar{Y}) = 0$$

has a stabilizing solution $\bar{Y}$, making $A_L := A - (D_w B_w' + \bar{Y} C')(D_w D_w')^{-1} C$ Hurwitz and such that $\bar{Y} = \bar{Y}' \geq 0$.

Remark. This CARE is dual to what we studied in the LQR case, just substitute

$$A \rightarrow A', \quad B \rightarrow C', \quad \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} B_w \\ D_w \end{bmatrix} \begin{bmatrix} B_w' & D_w' \end{bmatrix},$$

and is solved by `icare(A', C', Bw*Bw', Dw*Dw', Bw*Dw')`.

14/48

Problem solution

Theorem 9.1

If **A₁₋₃** hold, then the unique optimal linear-quadratic estimator is

$$\begin{aligned} \hat{\dot{x}}(t) = & A \hat{x}(t) + B_u u(t) \\ & + (B_w D_w' + \bar{Y} C')(D_w D_w')^{-1} (y(t) - C \hat{x}(t)), \quad \hat{x}(0) = 0 \end{aligned}$$

where $\bar{Y} = \bar{Y}' \geq 0$ is the stabilizing solution to the CARE

$$\bar{Y} A' + A \bar{Y} + B_w B_w' - (B_w D_w' + \bar{Y} C')(D_w D_w')^{-1} (D_w B_w' + C \bar{Y}) = 0.$$

The optimal cost $\mathcal{J}_{opt} = \operatorname{tr}(\bar{Y})$.

This solution is

- known as the (steady-state) **Kalman-Bucy filter** mainly under $u = 0$
- a full-order observer, with the gain $L = -(B_w D_w' + \bar{Y} C')(D_w D_w')^{-1}$

15/48

Proof outline

Assuming that the observer is $\dot{\hat{x}} = A \hat{x} + B_u u - L(y - C \hat{x})$, the error satisfies

$$\dot{\epsilon} = (A + LC)\epsilon + (B_w + LD_w)w.$$

Hence,

$$\epsilon_i(t) = e^{(A+LC)t} (B_w + LD_w) e_i$$

Because $a = \operatorname{tr}(a)$ and $\operatorname{tr}(AB) = \operatorname{tr}(BA)$,

$$\begin{aligned} \int_0^\infty \|\epsilon_i(t)\|^2 dt &= \int_0^\infty \operatorname{tr}(\epsilon_i'(t) \epsilon_i(t)) dt = \int_0^\infty \operatorname{tr}(\epsilon_i(t) \epsilon_i'(t)) dt \\ &= \int_0^\infty \operatorname{tr}(e^{(A+LC)t} (B_w + LD_w) e_i e_i' (B_w + LD_w)' e^{(A+LC)'t}) dt \\ &= \operatorname{tr}(W_i), \end{aligned}$$

where $W_i = W_i' \geq 0$ solves the Lyapunov equation

$$(A + LC)W_i + W_i(A + LC)' + (B_w + LD_w) e_i e_i' (B_w + LD_w)' = 0.$$

16/48

Proof outline (contd)

Thus,

$$\mathcal{J} = \sum_{i=1}^{m_w} \int_0^\infty \|\epsilon_i(t)\|^2 dt = \text{tr} \left(\sum_{i=1}^{m_w} W_i \right) = \text{tr}(W),$$

where $W := \sum_{i=1}^{m_w} W_i \geq 0$ satisfies

$$(A + LC)W + W(A + LC)' + (B_w + LD_w) \left(\sum_{i=1}^{m_w} e_i e_i' \right) (B_w + LD_w)' = 0$$

or, taking into account that $\sum_{i=1}^{m_w} e_i e_i' = I_{m_w}$,

$$(A + LC)W + W(A + LC)' + (B_w + LD_w)(B_w + LD_w)' = 0.$$

The fact that the minimizing W solves the filtering CARE could be seen via arguments of Lecture 6, Slide 37 (alas, they were not detailed either). $\square$

17/48

State is a key

If the cost function

$$\mathcal{J} = \sum_{i=1}^{m_w} \int_0^\infty \|z(t) - \hat{z}_i(t)\|^2 dt$$

for $z = C_z x$ and a given $C_z \in \mathbb{R}^{p_z \times n}$, then the optimal reconstruction is

$$\hat{z} = C_z \hat{x},$$

where $\hat{x}$ is the optimal reconstruction of the state x as in Theorem 9.1. The optimal cost in this case does depend on the choice of z ,

$$\mathcal{J}_{\text{opt}} = \text{tr}(C_z \bar{Y} C_z').$$

18/48

Example 2

Consider the system (the double integrator, acceleration $\mapsto$ position)

$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (d + u) \\ y = \begin{bmatrix} 1 & 0 \end{bmatrix} x + n \end{cases}$$

with the goal to reconstruct the velocity $z = x_2 = v$. Assume that d and n are independent and white, with the intensities $\sigma_d > 0$ and $\sigma_n > 0$. Hence,

$$w = \overbrace{\begin{bmatrix} d/\sqrt{\sigma_d} \\ n/\sqrt{\sigma_n} \end{bmatrix}}^{\text{unit-intensity}} \Rightarrow \begin{bmatrix} A & B_w & B_u \\ C & D_w & \\ C_z & & \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{\sigma_d} & 0 & 1 \\ 1 & 0 & 0 & \sqrt{\sigma_n} & \\ 0 & 1 & & & \end{bmatrix}.$$

Worth emphasizing is that

- σ_d and σ_n are assumed intensities, used to tune the designed observer.

The actual disturbance and noise may be different, not even white.

19/48

Example 2: the observer

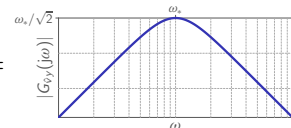
$\mathcal{A}_{1-3}$ hold and

$$\bar{Y} = \sqrt{\sigma_d \sigma_n} \begin{bmatrix} \sqrt[4]{4\sigma_n/\sigma_d} & 1 \\ 1 & \sqrt[4]{4\sigma_d/\sigma_n} \end{bmatrix} \Rightarrow \mathcal{J}_{\text{opt}} = \sqrt[4]{4\sigma_d^3 \sigma_n}.$$

Furthermore, $L = - \begin{bmatrix} \sqrt[4]{4\sigma_d/\sigma_n} & \sqrt{\sigma_d/\sigma_n} \end{bmatrix} = - \begin{bmatrix} \sqrt{2}\omega_* & \omega_*^2 \end{bmatrix}$ and

$$G_{\hat{v}y}(s) = \frac{\omega_*^2 s}{s^2 + \sqrt{2}\omega_* s + \omega_*^2}, \quad \text{where } \omega_* := \sqrt[4]{\frac{\sigma_d}{\sigma_n}}$$

whose frequency response magnitude

$$|G_{\hat{v}y}(j\omega)| = \frac{\omega}{\sqrt{1 + \omega^4/\omega_*^4}} =$$


Thus,

- larger $\sigma_d/\sigma_n \Rightarrow$ wider frequency band in which $G_{\hat{v}y}(j\omega) \approx j\omega$ (which corresponds to $v \approx \dot{y}$).

20/48

Example 2: SNR insight into the Kalman–Bucy filter

If $u = 0$, then the measurement

$$Y(s) = \frac{1}{s^2} D(s) + N(s) := \overbrace{Y_s(s)}^{\text{signal}} + \overbrace{Y_n(s)}^{\text{noise}}$$

Thus, it can be thought of as comprising

- the “signal” y_s , containing information about the signal to be estimated
- the “noise” y_n , containing no (useful) information about x

Convenient measure of the “usefulness” of y is the **signal to noise ratio**

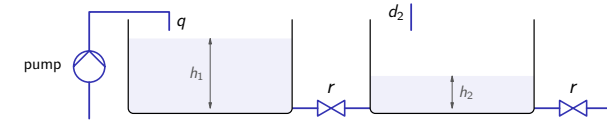
$$\text{SNR}(\omega) := \frac{|Y_s(j\omega)|}{|Y_n(j\omega)|} = \frac{\sqrt{\sigma_d}}{\sqrt{\sigma_n} \omega^2} = \frac{\omega_*^2}{\omega^2}.$$

At frequencies where $\text{SNR}(\omega) \gg 1$ measurements are informative, while at frequencies where $\text{SNR}(\omega) \ll 1$ we measure mostly noise. The filter then

- differentiates at frequencies where the assumed $\text{SNR}(\omega) \gg 1$
- ignores measurements at frequencies where the assumed $\text{SNR}(\omega) \ll 1$

21/48

Example: double-tank



Return to

$$\begin{cases} \dot{x}(t) = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} d_2(t), & x(0) = x_0 \\ y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) + n(t) \end{cases}$$

with

- measured (control) input $u(t) = 0.5 \sin(0.5\pi t) \mathbb{1}(t)$
- disturbance $d_2(t) = 0.05 \mathbb{1}(t)$
- measurement noise $n(t) = 0.01 \sin(4\pi t) \mathbb{1}(t)$

now with the linear-quadratic observer design.

22/48

Example: designs considerations

We already saw that only the ratio between intensities of d and n matters. So we take $\sigma_d = 1$ and tune σ_n . Its values are selected as

$$\sigma_n \in \{0.0356, 0.00042, 10^{-7}\},$$

tuned to have $T_{\epsilon_z d}(0)$ matching the pole-placement design of the full-order observer before. It results in

$$L = - \begin{bmatrix} 0.2206 & 0.2206 & 7.6571 \\ 0.4638 & 0.4638 & 300.8136 \end{bmatrix} \Rightarrow \text{spec}(A_L) = \begin{cases} -2.11 \pm j0.97 \\ -5.11 \pm j4.76 \\ -39.8 \pm j39.7 \end{cases}$$

Interesting is that the gains are substantially lower than the corresponding

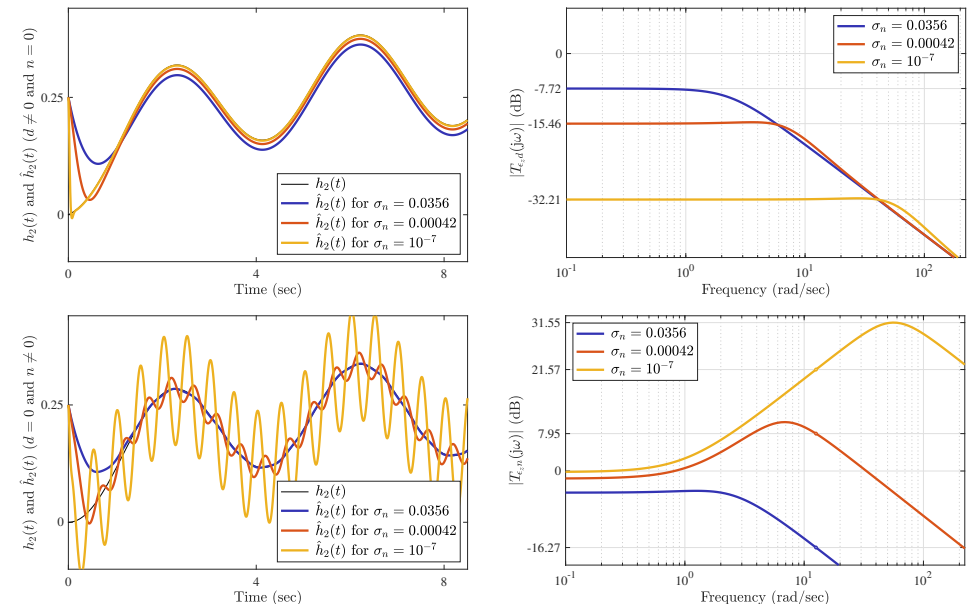
$$L_{\text{pole-placement}} = - \begin{bmatrix} -1.4 & 9.8 & 99.4 \\ 2.8 & 43.4 & 3896.2 \end{bmatrix}$$

(having the same $T_{\epsilon_z d}(0)$). As we can expect,

- smaller $\sigma_n \Rightarrow$ higher sensitivity to n and lower sensitivity to d .

23/48

Example: Kalman–Bucy filter simulations



24/48

Outline

Observers in the presence disturbances and measurement noise

Linear-quadratic estimation

Coloring linear-quadratic estimation

Linear Quadratic Gaussian (LQG) problem

LQG controller design (optional self-study)

25/48

Is white noise assumption realistic?

The white noise assumption about w of the least-square observer might not reflect properties of d and n accurately.

- Low-frequency harmonics typically dominate the spectrum of d . It may then be more natural to assume that

$$|D(j\omega)|^2 = \Phi_d(\omega)$$

for some $\Phi_d(\omega) \geq 0$ (power **spectral density**), large at low frequencies and small at high frequencies (like $a^2/(\omega^2 + a^2)$, $a \neq 0$).

- High-frequency harmonics typically dominate the spectrum of n . It may then be more natural to assume that

$$|N(j\omega)|^2 = \Phi_n(\omega)$$

for some $\Phi_n(\omega) \geq 0$ (power spectral density), small at low frequencies and large at high frequencies (like $\omega^2/(\omega^2 + a^2)$, $a \neq 0$).

Such information can be incorporated into the least-square framework.

26/48

Coloring transformation

The idea is to generate d and n as the impulse responses of LTI systems,

$$W_d : \begin{cases} \dot{x}_d(t) = A_d x_d(t) + B_d w_d(t), & x_d(0) = 0 \\ d(t) = C_d x_d(t) \end{cases}$$

(with a strictly proper $W_d(s)$, reflecting a low-frequency domination of the spectrum of d) and

$$W_n : \begin{cases} \dot{x}_n(t) = A_n x_n(t) + B_n w_n(t), & x_n(0) = 0 \\ n(t) = C_n x_n(t) + D_n w_n(t) \end{cases}$$

(a bi-proper $W_n(s)$ reflects a high-frequency domination of the spectrum of n) for white w_d and w_n . In this case

$$|D(j\omega)|^2 = |W_d(j\omega)|^2 \quad \text{and} \quad |N(j\omega)|^2 = |W_n(j\omega)|^2$$

are not constant. This is what known as **color noise**.

27/48

Coloring transformation (contd)

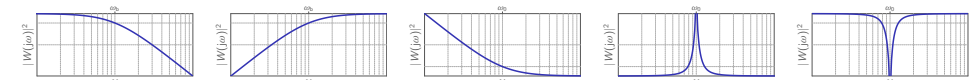
The technique is reminiscent of the use of exosystems, yet now with a

- frequency-domain motivation.

For that reason, coloring filters do not have to be unstable, they often have their poles in the CLHP $\{z \in \mathbb{C} \mid \text{Re } z \leq 0\}$. Some examples:

- low-pass filters, like Butterworth $W(s) = 1/B_n(s/\omega_b)$
- high-pass filters, like Butterworth $W(s) = (s/\omega_b)^n/B_n(s/\omega_b)$
- PI $W(s) = (s + \omega_0)/s$ overemphasizes the zero frequency
- harmonic oscillator $W(s) = (s^2 + 2\zeta\omega_0 s + \omega_0^2)/(s^2 + \omega_0^2)$
- notch, like $W(s) = (s^2 + \omega_0^2)/(s^2 + 2\zeta\omega_0 s + \omega_0^2)$

with their Bode magnitude plots



28/48

Example: Dryden gusts

A widely used model of the lateral wind turbulence for an aircraft flying at a (constant) speed v through a frozen turbulence field is $d = W_{\text{drdn}} w_d$, where

$$W_{\text{drdn}}(s) = \sigma_v \sqrt{\frac{L_v v}{\pi}} \frac{\sqrt{3} L_v s + v}{(L_v s + v)^2},$$

and σ_v and L_v represent turbulence intensity and scale length, respectively, and w_d is a white unit-intensity signal. In this case

$$|D(j\omega)|^2 = |W_{\text{drdn}}(j\omega)|^2 = \frac{\sigma_v^2 L_v v}{\pi} \frac{3L_v^2 \omega^2 + v^2}{(L_v^2 \omega^2 + v^2)^2} =$$


Similar models exist for longitudinal and vertical gusts.

29/48

State reconstruction with coloring transformation

If

$$\begin{cases} \dot{x} = Ax + B_d d + B u \\ y = Cx + n \end{cases}, \quad \begin{cases} \dot{x}_d = A_d x_d + B_d w_d \\ d = C_d x_d \end{cases}, \quad \begin{cases} \dot{x}_n = A_n x_n + B_n w_n \\ n = C_n x_n + D_n w_n \end{cases}$$

then the overall system can be described as

$$\begin{cases} \begin{bmatrix} \dot{x} \\ \dot{x}_d \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} A & B_d C_d & 0 \\ 0 & A_d & 0 \\ 0 & 0 & A_n \end{bmatrix} \begin{bmatrix} x \\ x_d \\ x_n \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ B_d & 0 \\ 0 & B_n \end{bmatrix} \begin{bmatrix} w_d \\ w_n \end{bmatrix} + \begin{bmatrix} B \\ 0 \\ 0 \end{bmatrix} u \\ y = \begin{bmatrix} C & 0 & C_n \end{bmatrix} \begin{bmatrix} x \\ x_d \\ x_n \end{bmatrix} + \begin{bmatrix} 0 & D_n \end{bmatrix} \begin{bmatrix} w_d \\ w_n \end{bmatrix} \end{cases}$$

which is the same general system, under the white $w = \begin{bmatrix} w_d \\ w_n \end{bmatrix}$. Note that

$$-D_w B'_w = \begin{bmatrix} 0 & D_n B_n \end{bmatrix} \neq 0 \text{ in general}$$

meaning that process and measurement noises are not independent now.

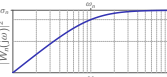
30/48

Example 2 with high-frequency noise

Return to the double integrator system, supposing now that

$$n = W_n w_n, \quad \text{where } W_n(s) = \frac{\sqrt{\sigma_n} s}{s + \omega_n} = \begin{bmatrix} -\omega_n & -\omega_n \\ \sqrt{\sigma_n} & \sqrt{\sigma_n} \end{bmatrix}.$$

This assumes that

$$|N(j\omega)|^2 = \frac{\omega^2}{\omega^2 + \omega_n^2} \sigma_n =$$


(i.e. that noise is mostly in high frequencies, $\omega > \omega_n$). In this case

$$w = \begin{bmatrix} \text{unit-intensity} \\ d/\sqrt{\sigma_d} \\ w_n \end{bmatrix} \Rightarrow \begin{bmatrix} A & B_w & B_u \\ C & D_w \\ C_z \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{\sigma_d} & 0 & 1 \\ 0 & 0 & -\omega_n & 0 & -\omega_n & 0 \\ 1 & 0 & \sqrt{\sigma_n} & 0 & \sqrt{\sigma_n} & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

with observable (C, A) .

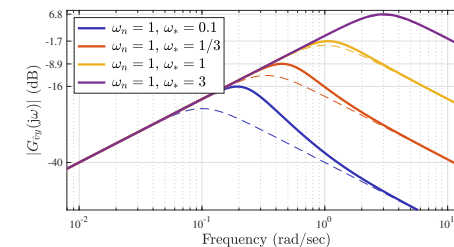
31/48

Example 2 with high-frequency noise (contd)

The measurements under $u = 0$ are now

$$Y(s) = \frac{1}{s^2} D(s) + \frac{\sqrt{\sigma_n} s}{s + \omega_n} \Rightarrow \text{SNR}(\omega) = \frac{\omega_*^2}{\omega^2} \sqrt{1 + \frac{\omega_n^2}{\omega^2}} > \frac{\omega_*^2}{\omega^2},$$

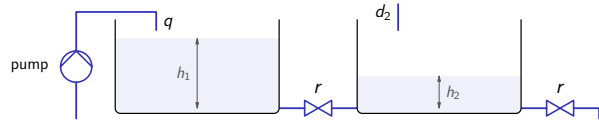
This results in



with a wider frequency band where $G_y(j\omega) \approx j\omega$ if $\frac{\omega_n}{\omega_*}$ is sufficiently large, i.e. if the increase of $\text{SNR}(\omega_*)$ is substantial.

32/48

Example: double-tank



Return to

$$\begin{cases} \dot{x}(t) = \begin{bmatrix} -1 & 1 \\ 1 & -2 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix} d_2(t), & x(0) = x_0 \\ y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} x(t) + n(t) \end{cases}$$

with

- measured (control) input $u(t) = 0.5 \sin(0.5\pi t) \mathbb{1}(t)$
- disturbance $d_2(t) = 0.05 \mathbb{1}(t)$
- measurement noise $n(t) = 0.01 \sin(4\pi t) \mathbb{1}(t)$

now with the linear-quadratic observer design.

33/48

Example: design considerations

Assuming a priori knowledge of dominant frequencies of d and n , choose

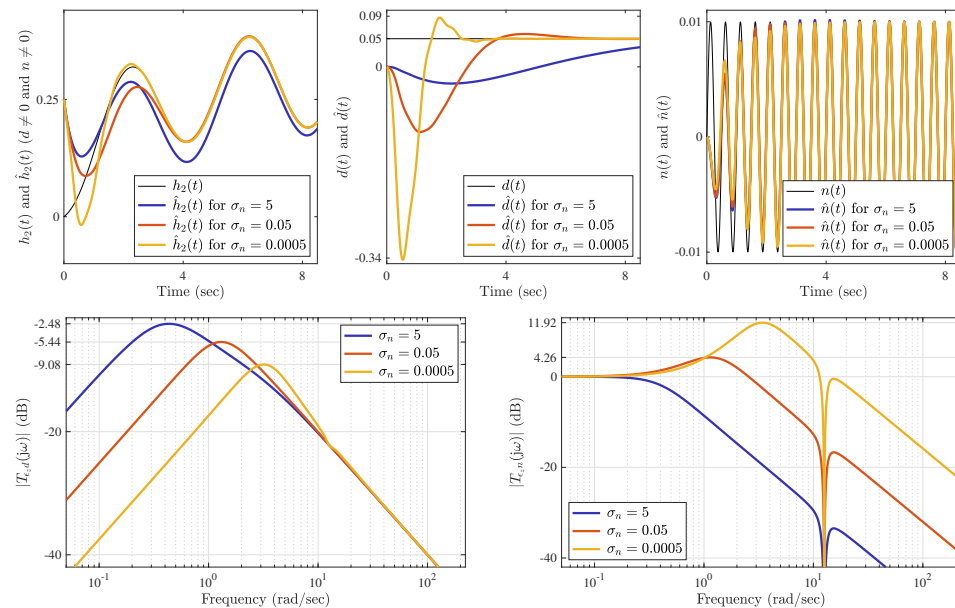
$$W_d(s) = \frac{1}{s} \quad \text{and} \quad W_n(s) = \sigma_n \frac{s^2 + 0.8\pi s + 16\pi^2}{s^2 + 16\pi^2}$$

(the damping ratio in the numerator of $W_n(s)$ could be tuned). This result in a 5-dimensional state to reconstruct, whose

- first two components are x_1 and x_2 , as before
- third component is d
- last two components are states of W_n , so n is their linear combination
specific combination depends on the realization chosen

34/48

Example: simulations with colored disturbance and noise



35/48

Outline

Observers in the presence disturbances and measurement noise

Linear-quadratic estimation

Coloring linear-quadratic estimation

Linear Quadratic Gaussian (LQG) problem

LQG controller design (optional self-study)

36/48

The standard problem

Consider

$$\begin{cases} \dot{x}(t) = Ax(t) + B_w w(t) + B_u u(t), & x(0) = 0 \\ z(t) = C_z x(t) + D_{zu} u(t) \\ y(t) = C_y x(t) + D_{yw} w(t) \end{cases}$$

where

w is an exogenous input z is a regulated output (to be kept small)

u is a control input y is a measured output

assuming

$\mathcal{A}_4$: (C_y, A, B_u) is stabilizable and detectable

$\mathcal{A}_5$: $D'_{zu} D_{zu} > 0$ and $D_{yw} D'_{yw} > 0$

$\mathcal{A}_6$: $\begin{bmatrix} A - \lambda I & B_u \\ C_z & D_{zu} \end{bmatrix}$ has full column rank for all $\lambda \in \mathbb{R}$

$\mathcal{A}_7$: $\begin{bmatrix} A - \lambda I & B_w \\ C_y & D_{yw} \end{bmatrix}$ has full row rank for all $\lambda \in \mathbb{R}$

37/48

LQG version of the standard problem

Let z_i denote the regulated output under $w = e_i \delta$. The cost function to be minimized in the LQG approach is

$$\mathcal{J} = \sum_{i=1}^{m_w} \int_0^\infty \|z_i(t)\|^2 dt$$

It can also be written as

$$\mathcal{J} = \sum_{i=1}^{m_w} \int_0^\infty \begin{bmatrix} x'_i(t) & u'_i(t) \end{bmatrix} \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \begin{bmatrix} x_i(t) \\ u_i(t) \end{bmatrix} dt,$$

where

$$\begin{bmatrix} Q & S \\ S' & R \end{bmatrix} = \begin{bmatrix} C'_z \\ D'_{zu} \end{bmatrix} \begin{bmatrix} C_z & D_{zu} \end{bmatrix}$$

and x_i and u_i are those generated under $w = e_i \delta$.

38/48

LQG problem: solution (the separation principle)

Theorem 9.2

If $\mathcal{A}_{4-7}$ hold, then the unique optimal controller is

$$\begin{cases} \dot{\hat{x}}(t) = (A + B_u K + L C_y) \hat{x}(t) - L y(t), & \hat{x}(0) = 0 \\ u(t) = K \hat{x}(t) \end{cases}$$

where

$$K = -(D'_{zu} D_{zu})^{-1} (D'_{zu} C_z + B'_u \bar{X}) \text{ and } L = -(B_w D'_{yw} + \bar{Y} C'_y) (D_{yw} D'_{yw})^{-1}$$

and $\bar{X} = \bar{X} \geq 0$ and $\bar{Y} = \bar{Y} \geq 0$ are the stabilizing solutions to the CAREs

$$\begin{aligned} A' \bar{X} + \bar{X} A + C'_z C_z - (C'_z D_{zu} + \bar{X} B_u) (D'_{zu} D_{zu})^{-1} (D'_{zu} C_z + B'_u \bar{X}) &= 0 \\ A \bar{Y} + \bar{Y} A' + B_w B'_w - (B_w D'_{yw} + \bar{Y} C'_y) (D_{yw} D'_{yw})^{-1} (D_{yw} B'_w + C_y \bar{Y}) &= 0 \end{aligned}$$

The optimal cost $\mathcal{J}_{opt} = \text{tr}(B'_w \bar{X} B_w) + \text{tr}(C'_z \bar{Y} C'_z) + \text{tr}(\bar{X} A \bar{Y} + \bar{Y} A' \bar{X})$ then.

39/48

LQG problem: solution properties

- the optimal solution is an observer-based controller, comprising
 - state-feedback LQR gain K
 - linear-quadratic (Kalman–Bucy) observer gain L
- the optimal controller is stabilizing, but not necessarily stable itself;
- stability margins of LQR are no longer guaranteed, even if $C'_z D_{zu} = 0$;
- because

$$\text{tr}(B'_w \bar{X} B_w) = \sum_{i=1}^{m_w} e'_i B'_w \bar{X} B_w e_i,$$

the first term of $\mathcal{J}_{opt}$ is the sum (average) of LQR costs for $x_0 = B_w e_i$;

- because $\text{tr}(AB) = \text{tr}(BA)$ and $\text{tr}(A) + \text{tr}(B) = \text{tr}(A + B)$,

$$\begin{aligned} \text{tr}(C'_z \bar{Y} C'_z) + \text{tr}(\bar{X} A \bar{Y} + \bar{Y} A' \bar{X}) &= \text{tr}(C'_z C_z \bar{Y} + \bar{X} A \bar{Y} + A' \bar{X} \bar{Y}) \\ &= \text{tr}(K' D'_{zu} D_{zu} K \bar{Y}) \\ &= \text{tr}(D_{zu} K \bar{Y} K' D'_{zu}), \end{aligned}$$

the other terms of $\mathcal{J}_{opt}$ are the cost of reconstructing $D_{zu} K x$ (cf. z).

40/48

Outline

Observers in the presence disturbances and measurement noise

Linear-quadratic estimation

Coloring linear-quadratic estimation

Linear Quadratic Gaussian (LQG) problem

LQG controller design (optional self-study)

41/48

Think loop shaping, act LQG

Loop shaping:

- easy on the magnitude shaping
- might be intricate in the phase shaping around crossover especially in the MIMO case

LQG:

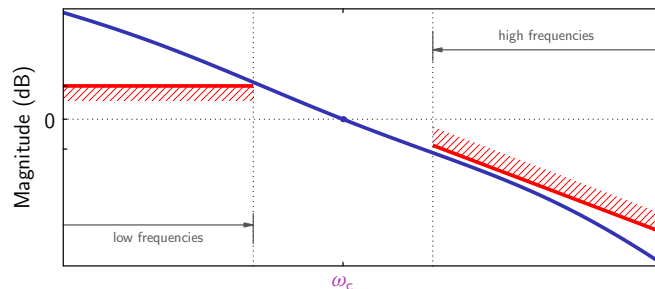
- guarantees closed-loop stability
- not directly interpretable in terms of frequency-domain properties

Combine:

- shape loop magnitude (via frequency-dependent weights)
- solve some LQG to stabilize the shaped loop

42/48

Magnitude shaping



Conceptually simple,

- pick W such that $|P_a(j\omega)|$, where $P_a = WP$, is shaped appropriately
 - possible choices
 - $W(s) = k_p$ to fix the crossover
 - $W(s) = k_p \omega_c / s$ (I) or $W(s) = k_p(1 + k_i/s)$ (PI) to add integral action
 - $W(s) = k_p / (\tau s + 1)$ to add low-pass filter
- whatever R_a designed for P_a is then implemented as $R = R_a W$ for P

43/48

LQG

Assuming load disturbances, bring in

$$P_a : \begin{cases} \dot{x}_a(t) = A_a x_a(t) + B_a u(t) + B_a d(t), & x_a(0) = 0, \\ y_a(t) = C_a x_a(t) + n_a(t), \end{cases}$$

and consider a **balanced** LQG for it with the cost

$$\mathcal{J} = \int_0^\infty (y_a^2(t) + u^2(t)) dt = \int_0^\infty \begin{bmatrix} x_a'(t) & u'(t) \end{bmatrix} \begin{bmatrix} C_a' C_a & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_a(t) \\ u(t) \end{bmatrix} dt$$

and unit intensities of d and n_a . In other words,

- cost puts equal importance to $y_a = Wy$ and u ,
- load disturbance d and noise $n_a = Wn$ are equally important.

Motivation:

- LQG shall just ensure stability $\implies$ unsophisticated and balanced.

44/48

Controller

The resulted controller is

$$R(s) = R_a(s)W(s).$$

Some of its properties:

- the degree of $R(s)$, $\deg(R(s)) = \deg(P(s)) + 2\deg(W(s))$ unless some poles or zeros of $W(s)$ are canceled by $R_a(s)$
- integrators of $W(s)$ become integrators of $R(s)$
- any nonzero roll-off in $W(s)$ increases that of $R(s)$

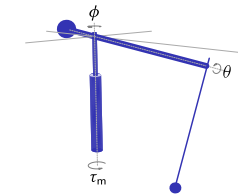
Success indicator:

$$\epsilon = \frac{1}{\sqrt{1 + \lambda_{\max}(\bar{X}_a \bar{Y}_a)}} < 1,$$

where $\bar{X}_a \geq 0$ and $\bar{Y}_a \geq 0$ are the stabilizing solutions to the CAREs for P_a . The level $\epsilon \geq 0.25$ may be regarded as adequate, $\epsilon \geq 0.5$ is very good.

45/48

Exercise: Furuta pendulum dampening



whose model is 4-order, with the state vector $x = (\phi, \dot{\phi}, \theta, \dot{\theta})$ and

$$\begin{bmatrix} A & B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & J_b + mr^2 & 0 & mrl \\ 0 & 0 & 1 & 0 \\ 0 & mrl & 0 & J_p + ml^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -gml & -\zeta \end{bmatrix}.$$

and

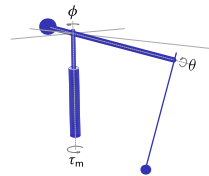
$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

46/48

Exercise: Furuta pendulum dampening (contd)

The control goal is twofold,

1. set-point tracking for the arm
i.e. moving ϕ from $\phi(0) = 0$ to some $\phi(t_f) = \phi_f$
2. dampening the pendulum
i.e. ensure $|T_d(j\omega)| \ll |P_{\phi\tau_m}(j\omega)|$ for ω 's around the resonance



We have one control input, τ_m , and two regulated outputs, ϕ and θ . Such systems are dubbed *underactuated*. Lucky we,

- control goals for ϕ and θ are often separated in the frequency domain, because
- dampening is only concerned with frequencies around the resonance
- set-point tracking is a low-frequency goal unless we require too fast settling

47/48

Exercise: Furuta pendulum dampening

- may start with $W(s) = \begin{bmatrix} k_\phi & 0 \\ 0 & k_\theta \end{bmatrix}$
- $W(s) = \begin{bmatrix} k_\phi(1 + b_\phi/s) & 0 \\ 0 & k_\theta \end{bmatrix}$ adds integral action to the ϕ -channel
- $W(s) = \begin{bmatrix} k_\phi & 0 \\ 0 & k_\theta((s + b_\theta)) \end{bmatrix}$ adds phase lead to the θ -channel

48/48