Balanced Truncation

Charbel Farhat
Stanford University
cfarhat@stanford.edu

These slides are based on the recommended textbook: A.C. Antoulas, “Approximation of Large-Scale Dynamical Systems,” Advances in Design and Control, SIAM, ISBN-0-89871-529-6
Outline

1. Reachability and Observability
2. Balancing
3. Balanced Truncation Method
4. Error Analysis
5. Stability Analysis
6. Computational Complexity
7. Comparison with the POD Method
8. Application
9. Balanced POD Method
Consider the following **stable**, high-dimensional, LTI system

\[
\begin{align*}
\frac{dw}{dt}(t) &= Aw(t) + Bu(t) \\
y(t) &= Cw(t) \\
w(0) &= w_0
\end{align*}
\]

- \(w \in \mathbb{R}^N\): State variables
- \(u \in \mathbb{R}^p\): Input variables, typically \(p \ll N\)
- \(y \in \mathbb{R}^q\): Output variables, typically \(q \ll N\)

Recall that the solution \(w(t)\) of the above linear ODE can be written as

\[
w(t) = \phi(t, u; t_0, w_0) = e^{A(t-t_0)}w(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau, \quad \forall t \geq t_0
\]
Reachability and Observability

Reachability, Controllability, and Observability

Definition (1)

A state \( w \in \mathbb{R}^N \) is **reachable** if there exists an input function \( u(.) \) of finite energy and a time \( T < \infty \) such that under this input and zero initial condition, the state of the system is \( w \).
Reachability and Observability

Reachability, Controllability, and Observability

**Definition (1)**

A state \( w \in \mathbb{R}^N \) is **reachable** if there exists an input function \( u(.) \) of finite energy and a time \( T < \infty \) such that under this input and zero initial condition, the state of the system is \( w \)

**Definition (2)**

A state \( w \in \mathbb{R}^N \) is **controllable** to the zero state if there exist an input function \( u(.) \) and a time \( T < \infty \) such that

\[
\phi(T, u; 0, w) = 0_N
\]
Reachability and Observability

Reachability, Controllability, and Observability

**Definition (1)**
A state $w \in \mathbb{R}^N$ is **reachable** if there exists an input function $u(.)$ of finite energy and a time $T < \infty$ such that under this input and zero initial condition, the state of the system is $w$.

**Definition (2)**
A state $w \in \mathbb{R}^N$ is **controllable** to the zero state if there exist an input function $u(.)$ and a time $T < \infty$ such that

$$
\phi(T, u; 0, w) = 0_N
$$

**Definition (3)**
A state $w \in \mathbb{R}^N$ is **unobservable** if for all $t \geq 0$,

$$
y(t) = C\phi(t, 0; 0, w) = 0_q
$$
A linear dynamical system \((A, B, C)\) is **completely controllable** at time \(t_0\) if it is not equivalent, for all \(t \geq t_0\), to a system of the type

\[
\begin{align*}
\frac{dw^{(1)}}{dt} &= A^{(1,1)}w^{(1)} + A^{(1,2)}w^{(2)} + B^{(1)}u \\
\frac{dw^{(2)}}{dt} &= A^{(2,2)}w^{(2)} \\
y(t) &= C^{(1)}w^{(1)} + C^{(2)}w^{(2)}
\end{align*}
\]

Interpretation: It is not possible to find a coordinate system in which the state variables are separated into two groups, \(w^{(1)}\) and \(w^{(2)}\), such that the second group is affected neither by the first group, nor by the inputs to the system.

Controllability is only a property of \((A, B)\).

This definition can be extended to linear time-variant systems.
A linear dynamical system \((A, B, C)\) is **completely controllable** at time \(t_0\) if it is not equivalent, for all \(t \geq t_0\), to a system of the type

\[
\begin{align*}
\frac{d\mathbf{w}^{(1)}}{dt} &= A^{(1,1)}\mathbf{w}^{(1)} + A^{(1,2)}\mathbf{w}^{(2)} + B^{(1)}u \\
\frac{d\mathbf{w}^{(2)}}{dt} &= A^{(2,2)}\mathbf{w}^{(2)} \\
y(t) &= C^{(1)}\mathbf{w}^{(1)} + C^{(2)}\mathbf{w}^{(2)}
\end{align*}
\]

**Interpretation:** It is not possible to find a coordinate system in which the state variables are separated into two groups, \(\mathbf{w}^{(1)}\) and \(\mathbf{w}^{(2)}\), such that the second group is affected neither by the first group, nor by the inputs to the system.
A linear dynamical system \((A, B, C)\) is \textbf{completely controllable} at time \(t_0\) if it is not equivalent, for all \(t \geq t_0\), to a system of the type

\[
\begin{align*}
\frac{d\mathbf{w}^{(1)}}{dt} &= A^{(1,1)} \mathbf{w}^{(1)} + A^{(1,2)} \mathbf{w}^{(2)} + B^{(1)} u \\
\frac{d\mathbf{w}^{(2)}}{dt} &= A^{(2,2)} \mathbf{w}^{(2)} \\
y(t) &= C^{(1)} \mathbf{w}^{(1)} + C^{(2)} \mathbf{w}^{(2)}
\end{align*}
\]

**Interpretation:** It is not possible to find a coordinate system in which the state variables are separated into two groups, \(\mathbf{w}^{(1)}\) and \(\mathbf{w}^{(2)}\), such that the second group is affected neither by the first group, nor by the inputs to the system.

**Controllability** is only a property of \((A, B)\).
Reachability and Observability

Completely Controllable Dynamical System

Definition (4, - R.E. Kalman, 1963)

A linear dynamical system \((A, B, C)\) is **completely controllable** at time \(t_0\) if it is not equivalent, for all \(t \geq t_0\), to a system of the type

\[
\begin{align*}
\frac{dw^{(1)}}{dt} &= A^{(1,1)}w^{(1)} + A^{(1,2)}w^{(2)} + B^{(1)}u \\
\frac{dw^{(2)}}{dt} &= A^{(2,2)}w^{(2)} \\
y(t) &= C^{(1)}w^{(1)} + C^{(2)}w^{(2)}
\end{align*}
\]

**Interpretation:** It is not possible to find a coordinate system in which the state variables are separated into two groups, \(w^{(1)}\) and \(w^{(2)}\), such that the second group is affected neither by the first group, nor by the inputs to the system.

- Controllability is only a property of \((A, B)\)
- This definition can be extended to linear time-variant systems
A linear dynamical system \((A, B, C)\) is **completely observable** at time \(t_0\) if it is not equivalent, for all \(t \leq t_0\), to any system of the type

\[
\begin{align*}
\frac{dw^{(1)}}{dt} &= A^{(1,1)}w^{(1)} + B^{(1)}u \\
\frac{dw^{(2)}}{dt} &= A^{(2,1)}w^{(1)} + A^{(2,2)}w^{(2)} + B^{(2)}u \\
y(t) &= C^{(1)}w^{(1)}
\end{align*}
\]
A linear dynamical system \((A, B, C)\) is **completely observable** at time \(t_0\) if it is not equivalent, for all \(t \leq t_0\), to any system of the type

\[
\begin{align*}
\frac{dw^{(1)}}{dt} &= A^{(1,1)} w^{(1)} + B^{(1)} u \\
\frac{dw^{(2)}}{dt} &= A^{(2,1)} w^{(1)} + A^{(2,2)} w^{(2)} + B^{(2)} u \\
y(t) &= C^{(1)} w^{(1)}
\end{align*}
\]

**Dual of the previous definition**
A linear dynamical system \((A, B, C)\) is **completely observable** at time \(t_0\) if it is not equivalent, for all \(t \leq t_0\), to any system of the type

\[
\begin{align*}
\frac{dw^{(1)}}{dt} &= A^{(1,1)}w^{(1)} + B^{(1)}u \\
\frac{dw^{(2)}}{dt} &= A^{(2,1)}w^{(1)} + A^{(2,2)}w^{(2)} + B^{(2)}u \\
y(t) &= C^{(1)}w^{(1)}
\end{align*}
\]

- Dual of the previous definition
- Interpretation: It is not possible to find a coordinate system in which the state variables are separated into two groups, \(w^{(1)}\) and \(w^{(2)}\), such that the second group affects neither the first group, nor the outputs of the system.
A linear dynamical system \((A, B, C)\) is **completely observable** at time \(t_0\) if it is not equivalent, for all \(t \leq t_0\), to any system of the type

\[
\begin{align*}
\frac{dw^{(1)}}{dt} &= A^{(1,1)}w^{(1)} + B^{(1)}u \\
\frac{dw^{(2)}}{dt} &= A^{(2,1)}w^{(1)} + A^{(2,2)}w^{(2)} + B^{(2)}u \\
y(t) &= C^{(1)}w^{(1)}
\end{align*}
\]

- Dual of the previous definition
- Interpretation: It is not possible to find a coordinate system in which the state variables are separated into two groups, \(w^{(1)}\) and \(w^{(2)}\), such that the second group affects neither the first group, nor the outputs of the system.
- Observability is only a property of \((A, C)\)
A linear dynamical system \((A, B, C)\) is **completely observable** at time \(t_0\) if it is not equivalent, for all \(t \leq t_0\), to any system of the type

\[
\begin{align*}
\frac{dw^{(1)}}{dt} &= A^{(1,1)}w^{(1)} + B^{(1)}u \\
\frac{dw^{(2)}}{dt} &= A^{(2,1)}w^{(1)} + A^{(2,2)}w^{(2)} + B^{(2)}u \\
y(t) &= C^{(1)}w^{(1)}
\end{align*}
\]

- Dual of the previous definition
- Interpretation: It is not possible to find a coordinate system in which the state variables are separated into two groups, \(w^{(1)}\) and \(w^{(2)}\), such that the second group affects neither the first group, nor the outputs of the system
- Observability is only a property of \((A, C)\)
- This definition can be extended to linear time-variant systems
The equation of the network in terms of the current $w_1$ flowing through the inductor and the potential $w_2$ across the capacitor ($LC = R = 1$) is given by

\[
\frac{dw_1}{dt} = -\frac{1}{L}w_1 + u_1
\]
\[
\frac{dw_2}{dt} = -\frac{1}{C}w_2 + u_1
\]
\[
y_1 = \frac{1}{L}w_1 + \frac{1}{C}w_2 + u_1
\]
Under the change of variable $\tilde{w}_1 = (w_1 + w_2)/2$ and $\tilde{w}_2 = (w_1 - w_2)/2$, the previous dynamical system becomes

\[
\frac{d\tilde{w}_1}{dt} = -\frac{1}{L} \tilde{w}_1 + u_1 \\
\frac{d\tilde{w}_2}{dt} = -\frac{1}{L} \tilde{w}_2 \\
y_1 = \frac{2}{L} \tilde{w}_2 + u_1
\]
Under the change of variable $\tilde{w}_1 = (w_1 + w_2)/2$ and $\tilde{w}_2 = (w_1 - w_2)/2$, the previous dynamical system becomes

$$\frac{d \tilde{w}_1}{dt} = -\frac{1}{L} \tilde{w}_1 + u_1$$

$$\frac{d \tilde{w}_2}{dt} = -\frac{1}{L} \tilde{w}_2$$

$$y_1 = \frac{2}{L} \tilde{w}_2 + u_1$$

$\tilde{w}_1$ is controllable but not observable
Under the change of variable \( \tilde{w}_1 = (w_1 + w_2)/2 \) and \( \tilde{w}_2 = (w_1 - w_2)/2 \), the previous dynamical system becomes

\[
\frac{d\tilde{w}_1}{dt} = -\frac{1}{L} \tilde{w}_1 + u_1 \\
\frac{d\tilde{w}_2}{dt} = -\frac{1}{L} \tilde{w}_2 \\
y_1 = \frac{2}{L} \tilde{w}_2 + u_1
\]

- \( \tilde{w}_1 \) is controllable but not observable
- \( \tilde{w}_2 \) is observable but not controllable
Theorem (Kalman 1961)

Consider a dynamical system \((A, B, C)\). Then:

(i) There is a state space coordinate system in which the components of the state vector can be decomposed into four parts

\[
\mathbf{w} = [w^{(a)} \ w^{(b)} \ w^{(c)} \ w^{(d)}]^T
\]

(ii) The sizes \(N_a, N_b, N_c\) and \(N_d\) of these vectors do not depend on the choice of basis

(iii) The system matrices take the form

\[
A = \begin{bmatrix}
A^{(a,a)} & A^{(a,b)} & A^{(a,c)} & A^{(a,d)} \\
0 & A^{(b,b)} & 0 & A^{(b,d)} \\
0 & 0 & A^{(c,c)} & A^{(c,d)} \\
0 & 0 & 0 & A^{(d,d)}
\end{bmatrix}, \quad B = \begin{bmatrix}
B^{(a)} \\
B^{(b)} \\
0 \\
0
\end{bmatrix},
\]

\[
C = \begin{bmatrix}
0 & C^{(b)} & 0 & C^{(d)}
\end{bmatrix}
\]
The four parts of \( w \) can be interpreted as follows
The four parts of $w$ can be interpreted as follows
- part (a) is completely controllable but unobservable
The four parts of \( w \) can be interpreted as follows:

- part (a) is completely controllable but unobservable
- part (b) is completely controllable and completely observable
The four parts of $w$ can be interpreted as follows:

- Part (a) is completely controllable but unobservable.
- Part (b) is completely controllable and completely observable.
- Part (c) is uncontrollable and unobservable.
The four parts of \( \mathbf{w} \) can be interpreted as follows:

- Part (a) is completely controllable but unobservable.
- Part (b) is completely controllable and completely observable.
- Part (c) is uncontrollable and unobservable.
- Part (d) is uncontrollable and completely observable.
The four parts of \( w \) can be interpreted as follows

- part (a) is completely controllable but unobservable
- part (b) is completely controllable and completely observable
- part (c) is uncontrollable and unobservable
- part (d) is uncontrollable and completely observable

This theorem can be extended to linear time-variant systems
Definition (6)

The **reachable subspace** $W_{\text{reach}} \subset \mathbb{R}^N$ of a system $(A, B, C)$ is the set containing all reachable states of the system.

$$R(A \ B) = [B \ AB \ \cdots \ A^{N-1}B \ \cdots]$$

is the **reachability matrix** of the system.
Definition (6)

The **reachable subspace** \( W_{reach} \subset \mathbb{R}^N \) of a system \((A, B, C)\) is the set containing all reachable states of the system.

\[
\mathcal{R}(A B) = [B \ AB \ \cdots \ A^{N-1}B \ \cdots]
\]

is the **reachability matrix** of the system.

Definition (7)

The **controllable subspace** \( W_{contr} \subset \mathbb{R}^N \) of a system \((A, B, C)\) is the set containing all controllable states of the system.
Theorem

Consider the system \((A, B, C)\). Then

\[ \mathcal{W}_{reach} = \text{Im} \mathcal{R}(A, B) \]
Theorem

Consider the system \((A, B, C)\). Then

\[ W_{\text{reach}} = \text{Im} \mathcal{R}(A, B) \]

Corollary

(i) \( A W_{\text{reach}} \subset W_{\text{reach}} \)
(ii) The system is completely reachable if and only if \( \text{rank} \mathcal{R}(A, B) = N \)
(iii) Reachability is basis independent
Reachability and Observability Gramians

**Definition**

The **reachability (controllability) Gramian** at time $t < \infty$ is defined as the $N \times N$ symmetric positive semi-definite matrix

$$\mathcal{P}(t) = \int_0^t e^{A\tau} BB^* e^{A^*\tau} d\tau$$
The **reachability (controllability) Gramian** at time $t < \infty$ is defined as the $N \times N$ symmetric positive semi-definite matrix

$$\mathcal{P}(t) = \int_0^t e^{A\tau} BB^* e^{A^*\tau} \, d\tau$$

The **observability Gramian** at time $t < \infty$ is defined as the $N \times N$ symmetric positive semi-definite matrix

$$\mathcal{Q}(t) = \int_0^t e^{A^*\tau} C^* C e^{A\tau} \, d\tau$$
Proposition

*The columns of $\mathcal{P}(t)$ span the reachability subspace*

$\mathcal{W}_{\text{reach}} = \text{Im } \mathcal{R}(A, B)$
Proposition

The columns of $\mathcal{P}(t)$ span the reachability subspace

$\mathcal{W}_{\text{reach}} = \text{Im } \mathcal{R}(A, B)$

Corollary

A system $(A, B, C)$ is reachable if and only if $\mathcal{P}(t)$ is symmetric positive definite at some time $t > 0$
Reachability and Observability

Equivalence Between Reachability and Controllability

Theorem

For continuous linear dynamical systems, the notions of controllability and reachability are equivalent – that is,

\[ W_{reach} = W_{contr} \]
Definition

The \textbf{unobservability subspace} \( \mathcal{W}_{\text{unobs}} \subset \mathbb{R}^N \) is the set of all unobservable states of the system. The matrix

\[
\mathcal{O}(C, A) = [C^* \ A^*C^* \ \cdots \ (A^*)^iC^* \ \cdots]^*
\]

is the \textbf{observability matrix} of the system.
Consider the system \((A, B, C)\). Then

\[
\mathcal{W}_{unobs} = \text{Ker } \mathcal{O}(C, A)
\]
Theorem

Consider the system \((A, B, C)\). Then

\[ W_{unobs} = \ker O(C, A) \]

Corollary

(i) \( A W_{unobs} \subset W_{unobs} \)

(ii) The system is completely observable if and only if \( \text{rank } O(C, A) = N \)

(iii) Observability is basis independent
Definition

The **infinite reachability (controllability) Gramian** is defined for a stable LTI system as the $N \times N$ symmetric positive semi-definite matrix

$$P = \int_0^\infty e^{At}BB^*e^{A^*t}dt$$
Definition

The infinite reachability (controllability) Gramian \( P \) is defined for a stable LTI system as the \( N \times N \) symmetric positive semi-definite matrix

\[
P = \int_0^\infty e^{At}BB^*e^{A^*t}dt
\]

Definition

The infinite observability Gramian \( Q \) is defined for a stable LTI system as the \( N \times N \) symmetric positive semi-definite matrix

\[
Q = \int_0^\infty e^{A^*t}C^*Ce^{At}dt
\]
Using Parseval’s theorem, the two previously defined Gramians can be written in the frequency domain as follows:
Using Parseval’s theorem, the two previously defined Gramians can be written in the frequency domain as follows:

**Infinite reachability Gramian**

\[
\mathcal{P} = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\omega I_N - A)^{-1}BB^*(-j\omega I_N - A^*)^{-1} d\omega
\]
Using Parseval’s theorem, the two previously defined Gramians can be written in the frequency domain as follows:

- **infinite reachability Gramian**

\[
P = \frac{1}{2\pi} \int_{-\infty}^{\infty} (j\omega I_N - A)^{-1} B B^* (-j\omega I_N - A^*)^{-1} d\omega
\]

- **infinite observability Gramian**

\[
Q = \frac{1}{2\pi} \int_{-\infty}^{\infty} (-j\omega I_N - A^*)^{-1} C^* C (j\omega I_N - A)^{-1} d\omega
\]
The two infinite Gramians are solutions of the following Lyapunov equations

- infinite reachability Gramian

\[
AP + PA^* + BB^* = 0_N
\]
The two infinite Gramians are solutions of the following Lyapunov equations

- infinite reachability Gramian

\[ \mathcal{A}P + P\mathcal{A}^* + BB^* = 0_N \]

- infinite observability Gramian

\[ A^*Q + QA + C^*C = 0_N \]
\( P \) and \( Q \) are respective bases for the reachable and observable subspaces
Reachability and Observability

Energetic Interpretation

- $\mathcal{P}$ and $\mathcal{Q}$ are respective bases for the reachable and observable subspaces
- $\| \cdot \|_{\mathcal{P}^{-1}}$ and $\| \cdot \|_{\mathcal{Q}}$ are semi-norms
\( \mathcal{P} \) and \( \mathcal{Q} \) are respective bases for the reachable and observable subspaces

- \( \| \| \mathcal{P}^{-1} \) and \( \| \| \mathcal{Q} \) are semi-norms

- For a reachable state, the inner product based on \( \mathcal{P}^{-1} \) characterizes the minimal energy required to steer the state from \( \mathbf{0} \) to \( \mathbf{w} \) as \( t \to \infty \)

\[
\| \mathbf{w} \|^2_{\mathcal{P}^{-1}} = \mathbf{w}^T \mathcal{P}^{-1} \mathbf{w} \quad \left( \leq \int_0^t (\mathbf{B}u(\tau))^* \mathbf{B}u(\tau) d\tau \right)
\]
$P$ and $Q$ are respective bases for the reachable and observable subspaces

- $\| \cdot \|_{P^{-1}}$ and $\| \cdot \|_{Q}$ are semi-norms

For a reachable state, the inner product based on $P^{-1}$ characterizes the minimal energy required to steer the state from $0$ to $w$ as $t \to \infty$

$$\|w\|_{P^{-1}}^2 = w^T P^{-1} w \quad \left( \leq \int_0^t (Bu(\tau))^* Bu(\tau) d\tau \right)$$

The inner product based on $Q$ indicates the maximal energy produced by observing the output of the system corresponding to an initial state $w_0$ when no input is applied

$$\|w\|_{Q}^2 = w^T Q w$$
PMOR strategy: **Eliminate** the states $w$ that are simultaneously
- PMOR strategy: **Eliminate** the states $\mathbf{w}$ that are simultaneously difficult to reach, i.e., require a large energy $\|\mathbf{w}\|_{P-1}^2$ to be reached.
PMOR strategy: **Eliminate** the states $w$ that are simultaneously

- difficult to reach, i.e., require a large energy $\|w\|_{P^{-1}}^2$ to be reached
- difficult to observe, i.e., produce a small observation energy $\|w\|_Q^2$
PMOR strategy: **Eliminate** the states $\mathbf{w}$ that are simultaneously
- **difficult to reach**, i.e., require a large energy $\|\mathbf{w}\|_{P^{-1}}^2$ to be reached
- **difficult to observe**, i.e., produce a small observation energy $\|\mathbf{w}\|_Q^2$

The above notions are **basis-dependent**
PMOR strategy: **Eliminate** the states $w$ that are simultaneously
- **difficult to reach**, i.e., require a large energy $\|w\|_{P^{-1}}^2$ to be reached
- **difficult to observe**, i.e., produce a small observation energy $\|w\|_Q^2$

The above notions are **basis-dependent**

One would like to consider a basis where both concepts are **equivalent**, i.e., where the system is **balanced**
Balancing requires changing the basis for the state using a transformation \( T \in \text{GL}(N) \)

\[ \hat{w} = Tw \]
Balancing requires changing the basis for the state using a transformation $T \in \text{GL}(N)$

$$\tilde{w} = Tw$$

Then
Balancing requires changing the basis for the state using a transformation \( T \in \text{GL}(N) \)

\[ \tilde{w} = Tw \]

Then

\[ e^{At}B \Rightarrow Te^{At}T^{-1}TB = Te^{At}B \]
Balancing requires changing the basis for the state using a transformation $T \in \text{GL}(N)$

$$\tilde{w} = Tw$$

Then

- $e^{At}B \Rightarrow Te^{At}T^{-1}TB = Te^{At}B$
- $B^*e^{A^*t} \Rightarrow B^*T^*T^{-*}e^{A^*t}T^* = B^*e^{A^*t}T^*$
Balancing requires changing the basis for the state using a transformation $T \in \text{GL}(N)$

$$\tilde{w} = Tw$$

Then

- $e^{At}B \Rightarrow Te^{At}T^{-1}TB = Te^{At}B$
- $B^*e^{A^*t} \Rightarrow B^*T^*T^{-1}e^{A^*t} = B^*e^{A^*t}T^*$
- the reachability Gramian becomes

$$\tilde{\mathcal{P}} = T\mathcal{P}T^*$$
Balancing requires changing the basis for the state using a transformation $T \in \text{GL}(N)$

$$\tilde{w} = Tw$$

Then

- $e^{At}B \Rightarrow Te^{At}T^{-1}TB = T e^{At}B$
- $B^* e^{A^*t} \Rightarrow B^* T^* T^{-*} e^{A^*t} T^* = B^* e^{A^*t} T^*$
- the reachability Gramian becomes

$$\tilde{P} = TP T^*$$

- $Ce^{At} \Rightarrow CT^{-1} Te^{At} T^{-1} = Ce^{At} T^{-1}$
Balancing requires changing the basis for the state using a transformation \( T \in \text{GL}(N) \)

\[
\tilde{w} = Tw
\]

Then

\[
e^{At}B \Rightarrow Te^{At}T^{-1}TB = Te^{At}B
\]

\[
B^*e^{A^*t} \Rightarrow B^*T^*T^{-*}e^{A^*t}T^* = B^*e^{A^*t}T^*
\]

the reachability Gramian becomes

\[
\tilde{\mathcal{P}} = T\mathcal{P}T^*
\]

\[
Ce^{At} \Rightarrow CT^{-1}Te^{At}T^{-1} = Ce^{At}T^{-1}
\]

\[
e^{A^*t}C^* \Rightarrow T^{-*}e^{A^*t}T^*T^{-*}C^* = T^{-*}e^{A^*t}C^*
\]
Balancing requires changing the basis for the state using a transformation $T \in \text{GL}(N)$

$$\tilde{w} = Tw$$

Then

- $e^{At}B \Rightarrow Te^{At}T^{-1}TB = Te^{At}B$
- $B^*e^{A^*t} \Rightarrow B^*TT^{-*}e^{A^*t}T^* = B^*e^{A^*t}T^*$
- the reachability Gramian becomes

$$\tilde{P} = TP^*T^*$$

- $Ce^{At} \Rightarrow CT^{-1}Te^{At}T^{-1} = Ce^{At}T^{-1}$
- $e^{A^*t}C^* \Rightarrow T^{-*}e^{A^*t}T^*T^{-*}C^* = T^{-*}e^{A^*t}C^*$
- the observability Gramian becomes

$$\tilde{Q} = T^{-*}QT^{-1}$$
The balancing transformations $T_{\text{bal}}$ and $T_{\text{bal}}^{-1}$ can be computed as follows:

1. Compute the Cholesky factorization $P = UU^*$
2. Compute the eigenvalue decomposition of $U^*QU$

$$U^*QU = K\Sigma^2K^*$$

where the entries in $\Sigma$ are ordered decreasingly.
3. Compute the transformations

$$T_{\text{bal}} = \Sigma^{1/2}K^*U^{-1}$$
$$T_{\text{bal}}^{-1} = UK\Sigma^{-1/2}$$
The balancing transformations \( T_{\text{bal}} \) and \( T_{\text{bal}}^{-1} \) can be computed as follows

1. compute the Cholesky factorization \( \mathcal{P} = \mathbf{U} \mathbf{U}^* \)
2. compute the eigenvalue decomposition of \( \mathbf{U}^* \mathbf{Q} \mathbf{U} \)
   \[ \mathbf{U}^* \mathbf{Q} \mathbf{U} = \mathbf{K} \mathbf{\Sigma}^2 \mathbf{K}^* \]
   where the entries in \( \mathbf{\Sigma} \) are ordered decreasingly
3. compute the transformations
   \[
   T_{\text{bal}} = \mathbf{\Sigma}^{1/2} \mathbf{K}^* \mathbf{U}^{-1}
   \]
   \[
   T_{\text{bal}}^{-1} = \mathbf{U} \mathbf{K} \mathbf{\Sigma}^{-1/2}
   \]

Then, one can check that balancing is achieved

\[
T_{\text{bal}} \mathcal{P} T_{\text{bal}}^* = T_{\text{bal}}^{-1} \mathbf{Q} T_{\text{bal}}^{-1} = \mathbf{\Sigma}
\]

**Definition (Hankel Singular Values)**

\( \mathbf{\Sigma} = \text{diag}(\sigma_1, \cdots, \sigma_N) \) contains the \( N \) Hankel singular values of the system
Computing the balancing transformation $T_{bal}$ is equivalent to minimizing the following function

$$\min_{T \in \text{GL}(N)} f(T) = \min_{T \in \text{GL}(N)} \text{trace}(TP^* + T^{-*}Q(T^{-1})$$
Computing the balancing transformation $T_{\text{bal}}$ is equivalent to minimizing the following function

$$\min_{T \in \text{GL}(N)} f(T) = \min_{T \in \text{GL}(N)} \text{trace}(TPT^* + T^{-*}QT^{-1})$$

For the optimal transformation $T_{\text{bal}}$, the function takes the value

$$f(T_{\text{bal}}) = 2\text{tr}(\Sigma) = 2 \sum_{i=1}^{N} \sigma_i$$

where $\{\sigma_i\}_{i=1}^{N}$ are the Hankel singular values.
Applying the change of variable $\tilde{w} = T_{\text{bal}}w$ transforms the given dynamical system into $(\tilde{A}, \tilde{B}, \tilde{C})$ where

$$T_{\text{bal}}AT_{\text{bal}}^{-1} = \tilde{A}, \quad T_{\text{bal}}B = \tilde{B}, \quad CT_{\text{bal}}^{-1} = \tilde{C}$$

Let $1 \leq k \leq N$; the system can be partitioned in blocks as

$$\tilde{A} = \begin{bmatrix} \tilde{A}_{11} & \tilde{A}_{12} \\ \tilde{A}_{21} & \tilde{A}_{22} \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} \tilde{B}_1 \\ \tilde{B}_2 \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} \tilde{C}_1 & \tilde{C}_2 \end{bmatrix}$$

The subscripts 1 and 2 denote the dimensions $k$ and $N - k$, respectively.
The blocks with the subscript 1 correspond to the most observable and reachable states.
The blocks with the subscript 1 correspond to the most observable and reachable states

The blocks with the subscript 2 correspond to the least observable and reachable states
Balanced Truncation Method

Block Partitioning of the System

- The blocks with the subscript 1 correspond to the most observable and reachable states
- The blocks with the subscript 2 correspond to the least observable and reachable states

Then, the following lower-dimensional model of size $k$

$$(A_r, B_r, C_r) = (\tilde{A}_{11}, \tilde{B}_1, \tilde{C}_1) \in \mathbb{R}^{k \times k} \times \mathbb{R}^{k \times p} \times \mathbb{R}^{q \times k}$$

is the PROM constructed by Balanced Truncation
The blocks with the subscript 1 correspond to the most observable and reachable states.

The blocks with the subscript 2 correspond to the least observable and reachable states.

Then, the following lower-dimensional model of size $k$

$$(A_r, B_r, C_r) = (\tilde{A}_{11}, \tilde{B}_1, \tilde{C}_1) \in \mathbb{R}^{k \times k} \times \mathbb{R}^{k \times p} \times \mathbb{R}^{q \times k}$$

is the PROM constructed by Balanced Truncation.

The left and right ROBs are

$$W = T^*_{bal}(:, 1 : k) \quad \text{and} \quad V = S_{bal}(:, 1 : k),$$

where $S_{bal} = T^{-1}_{bal}$.
Definition (The Hardy space $\mathcal{H}_\infty$)

The $\mathcal{H}_\infty$-norm associated with a system characterized by a transfer function $G(\cdot)$ is defined as

$$\|G\|_{\mathcal{H}_\infty} = \sup_{z \in \mathbb{C}_+} \|G(z)\|_{\infty} = \sup_{z \in \mathbb{C}_+} \sigma_{\text{max}}(G(z))$$

where $z \in \mathbb{C}_+$ if $z \in \mathbb{C}$ and $\text{Im}(z) > 0$.

Proposition

(i)  $\|G\|_{\mathcal{H}_\infty} = \sup_{\omega \in \mathbb{R}} \|G(i\omega)\|_{\infty} = \sup_{\omega \in \mathbb{R}} \sigma_{\text{max}}(G(i\omega))$

(ii) $\|G\|_{\mathcal{H}_\infty} = \sup_{u \neq 0} \frac{\|y(\cdot)\|_2}{\|u(\cdot)\|_2} = \sup_{u \neq 0} \sqrt{\frac{\int_0^\infty \|y(t)\|_2^2 dt}{\int_0^\infty \|u(t)\|_2^2 dt}}$

The $\mathcal{H}_\infty$ norm of the error between the HDM- and PROM-based solutions will be used as an error criterion.
Theorem (Error Bounds)

The Balanced Truncation procedure yields the following error bound for the output of interest. Let \( \{\tilde{\sigma}_i\}_{i=1}^{N_{SV}} \subseteq \{\sigma_i\}_{i=1}^{N} \) denote the distinct Hankel singular values of the system and \( \{\tilde{\sigma}_i\}_{i=N_k+1}^{N_{SV}} \) the ones that have been truncated. Then

\[
\|y(\cdot) - y_r(\cdot)\|_2 \leq 2 \sum_{i=N_k+1}^{N_{SV}} \tilde{\sigma}_i \|u(\cdot)\|_2
\]

Equivalently, the above result can be written in terms of the \( \mathcal{H}_\infty \)-norm of the system error as follows

\[
\|G(\cdot) - G_r(\cdot)\|_{\mathcal{H}_\infty} \leq 2 \sum_{i=N_k+1}^{N_{SV}} \tilde{\sigma}_i
\]

where \( G \) and \( G_r \) are the full- and reduced-order transfer functions. Equality holds when \( \tilde{\sigma}_{N_k+1} = \tilde{\sigma}_{N_{SV}} \) (all truncated singular values are equal).
Proof. The proof proceeds in 3 steps:

1. Consider the system error $E(s) = G(s) - G_r(s)$
Proof. The proof proceeds in 3 steps:

1. Consider the system error $E(s) = G(s) - G_r(s)$

2. Show that if all truncated singular values are equal to $\sigma$, then

$$\|E(\cdot)\|_{\mathcal{H}_\infty} = 2\sigma$$
Proof. The proof proceeds in 3 steps:

1. Consider the system error \( E(s) = G(s) - G_r(s) \)

2. Show that if all truncated singular values are equal to \( \sigma \), then

\[
\| E(\cdot) \|_{\mathcal{H}_\infty} = 2\sigma
\]

3. Use this result to show that in the general case

\[
\| E(\cdot) \|_{\mathcal{H}_\infty} \leq 2 \sum_{i=N_k+1}^{n_{SV}} \tilde{\sigma}_i
\]
Theorem (Stability Preservation)

Consider \((A_r, B_r, C_r) = (\tilde{A}_{11}, \tilde{B}_1, \tilde{C}_1)\), a PROM obtained by Balanced Truncation. Then

(i) \(A_r = \tilde{A}_1\) has no eigenvalues in the open right half plane

(ii) Furthermore, if the systems \((\tilde{A}_{11}, \tilde{B}_1, \tilde{C}_1)\) and \((\tilde{A}_{22}, \tilde{B}_2, \tilde{C}_2)\) have no Hankel singular values in common, \(A_r\) has no eigenvalues on the imaginary axis
Because of numerical stability issues, computing the transformations

\[ T_{\text{bal}} = \Sigma^{1/2} K^* U^{-1}, \quad T_{\text{bal}}^{-1} = UK\Sigma^{-1/2} \]

is usually ill-advised (computation of inverses)
Because of numerical stability issues, computing the transformations

\[ T_{\text{bal}} = \Sigma^{\frac{1}{2}} K^* U^{-1}, \quad T^{-1}_{\text{bal}} = U K \Sigma^{-\frac{1}{2}} \]

is usually ill-advised (computation of inverses).

Instead, it is better advised to use the following procedure:
Because of numerical stability issues, computing the transformations

\[ T_{bal} = \Sigma^{1/2} K^* U^{-1}, \quad T_{bal}^{-1} = UK\Sigma^{-1/2} \]

is usually ill-advised (computation of inverses)

Instead, it is better advised to use the following procedure

1. Start from the Cholesky decompositions of the Gramians

\[ P = UU^* \quad \text{and} \quad Q = ZZ^* \]
Because of numerical stability issues, computing the transformations

\[ T_{\text{bal}} = \Sigma^{1/2} K^* U^{-1}, \quad T_{\text{bal}}^{-1} = UK\Sigma^{-1/2} \]

is usually ill-advised (computation of inverses)

Instead, it is better advised to use the following procedure

1. start from the Cholesky decompositions of the Gramians
   \[ P = UU^* \quad \text{and} \quad Q = ZZ^* \]

2. compute the SVD
   \[ U^* Z = W\Sigma V^* \]
Because of numerical stability issues, computing the transformations

\[ T_{\text{bal}} = \Sigma^{\frac{1}{2}} K^* U^{-1} , \quad T_{\text{bal}}^{-1} = U K \Sigma^{-\frac{1}{2}} \]

is usually ill-advised (computation of inverses).

Instead, it is better advised to use the following procedure:

1. start from the Cholesky decompositions of the Gramians

   \[ P = U U^* \quad \text{and} \quad Q = Z Z^* \]

2. compute the SVD

   \[ U^* Z = W \Sigma V^* \]

3. construct the matrices

   \[ T_{\text{bal}} = \Sigma^{-\frac{1}{2}} V^* Z^* \quad \text{and} \quad T_{\text{bal}}^{-1} = U W \Sigma^{-\frac{1}{2}} \]
Balanced Truncation leads to PROMs with quality and stability guarantees; however
Balanced Truncation leads to PROMs with quality and stability guarantees; however

- the computation of a Gramian is intensive as it requires the solution of a Lyapunov equation ($O(N^3)$ operations)
Balanced Truncation leads to PROMs with quality and stability guarantees; however

- the computation of a Gramian is intensive as it requires the solution of a Lyapunov equation ($\mathcal{O}(N^3)$ operations)

- for this reason, Balanced Truncation is in general impractical for large systems – say $N \gtrsim 10^5$ (but monitor progress in the literature if interested)
Recall the theorem underlying the construction of a POD basis

**Theorem**

Let $\hat{K} \in \mathbb{R}^{N \times N}$ be the real symmetric semi-definite positive matrix defined as

$$\hat{K} = \int_{0}^{\mathcal{T}} w(t)w(t)^T dt$$

Let $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_N \geq 0$ denote the ordered eigenvalues of $\hat{K}$ and let $\hat{\phi}_i \in \mathbb{R}^N$, $i = 1, \cdots, N$, denote the associated eigenvectors

$$\hat{K}\hat{\phi}_i = \hat{\lambda}_i\hat{\phi}_i, \ i = 1, \cdots, N.$$ 

The subspace $\hat{V} = \text{range}(\hat{V})$ of dimension $k$ minimizing $J(\Pi_{V,V})$ is the invariant subspace of $\hat{K}$ associated with the eigenvalues $\hat{\lambda}_1, \cdots, \hat{\lambda}_k$. 
The response of an LTI system to a single impulse input with a zero initial condition is

\[ w(t) = e^{At}B \]
The response of an LTI system to a single impulse input with a zero initial condition is

\[ w(t) = e^{At}B \]

Consequently, the reachability Gramian is

\[ \hat{K} = \int_0^T w(t)w(t)^T dt = \int_0^T e^{At}BB^Te^{A^Tt} dt = \mathcal{P} \]
The response of an LTI system to a single impulse input with a zero initial condition is

$$w(t) = e^{At}B$$

Consequently, the reachability Gramian is

$$\hat{K} = \int_0^T w(t)w(t)^T dt = \int_0^T e^{At}BB^Te^{AT}t dt = \mathcal{P}$$

Unlike the Balanced Truncation method, the POD method does not take into account the observability Gramian to determine the PROM: therefore, every observable state may be truncated.
Objective: model the position of the lens controlled by a swing arm
System with $p = 2$ inputs and $q = 2$ outputs
Bode plots associated with the HDM-based solution ($N = 120$): Each column represents one input and each row represents a different output.
Bode plots associated with the PROM-based (Balanced Truncation) solution: Each column represents one input and each row represents a different output.
The Balanced POD Method generates two sets of snapshots: the standard POD solution snapshots; and the dual POD snapshots introduced below:

\[ S = [(j\omega_1I - A)^{-1}B \cdots (j\omega_{N_{\text{snap}}}I - A)^{-1}B] \]

\[ S_{\text{dual}} = [(-j\omega_1I - A^*)^{-1}C^* \cdots (-j\omega_{N_{\text{snap}}}I - A^*)^{-1}C^*] \]
The Balanced POD (BPOD) method generates two sets of snapshots: the standard POD solution snapshots; and the dual POD snapshots introduced below

\[
S = [(j\omega_1 I - A)^{-1} B \cdots (j\omega_{N\text{snap}} I - A)^{-1} B] \\
S_{\text{dual}} = [(-j\omega_1 I - A^*)^{-1} C^* \cdots (-j\omega_{N\text{snap}} I - A^*)^{-1} C^*]
\]

Next, BPOD computes right and left ROBs as follows

\[
S_{\text{dual}}^T S = U \Sigma Z^T \quad \text{(SVD)} \\
V = S Z_k \Sigma_k^{-1/2} \\
W = S_{\text{dual}} U_k \Sigma_k^{-1/2}
\]

where the subscript \( k \) designates the first \( k \) terms of the singular value decomposition.
The Balanced POD (BPOD) method generates two sets of snapshots: the standard POD solution snapshots; and the dual POD snapshots introduced below:

\[
S = [(j\omega_1 I - A)^{-1}B \cdots (j\omega_{\text{N}_{\text{snap}}} I - A)^{-1}B]
\]

\[
S_{\text{dual}} = [(-j\omega_1 I - A^*)^{-1}C^* \cdots (-j\omega_{\text{N}_{\text{snap}}} I - A^*)^{-1}C^*]
\]

Next, BPOD computes right and left ROBs as follows:

\[
S_{\text{dual}}^T S = U\Sigma Z^T \text{ (SVD)}
\]

\[
V = SZ_k \Sigma_k^{-1/2}
\]

\[
W = S_{\text{dual}}U_k \Sigma_k^{-1/2}
\]

where the subscript \( k \) designates the first \( k \) terms of the singular value decomposition.

If no truncation is performed, the result is equivalent to two-sided moment matching at \( s_i \in \{\omega_1, \cdots, \omega_{\text{N}_{\text{snap}}}\} \) (see later).
The POD method in the time domain is based solely on the reachability concept.
The POD method in the time domain is based solely on the reachability concept.

However, the BPOD method introduces also the notion of observability in the construction of a PROM. It is tractable for very large-scale systems and provides an approximation to the Balanced Truncation method. However, it does not guarantee in general the stability of the resulting PROM.
Supersonic Inlet Problem (part of the Oberwolfach Model Reduction Benchmark Collection repository)

\[ E \frac{d\mathbf{w}}{dt}(t) = A\mathbf{w}(t) + B\mathbf{u}(t) \]
\[ y(t) = C\mathbf{w}(t) \]

- \( N = 11,370 \) (2D Euler equations)
- \( p = 1 \) input (density disturbance of the inlet flow)
- \( q = 1 \) output (average Mach number at the diffuser throat)
PMOR in the frequency domain using
  - POD
  - BPOD
- In both cases, same frequency sampling for the computation of solution snapshots
Balanced POD Method

Application

- PMOR in the frequency domain using
  - POD
  - BPOD
- In both cases, same frequency sampling for the computation of solution snapshots
- Plot of the magnitude of the relative error in the transfer function (within the sampled frequency interval) as a function of the dimension $k$ of the constructed PROM