AuSRoS

Australian School of Robotic Systems

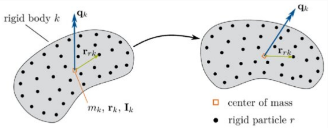
C2 Robotic state estimation

Guodong Shi

Rigid-body Robotics

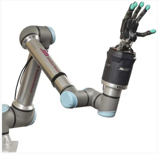
A rigid-body system is a collection of rigid bodies that may be connected together by joints and acted upon by various forces.

Rigid body assumption: In a body, the distance between any two points stays constant with respect to time.



Manipulator

Drone



Mobile robot

Bipedal robot "Cassie"

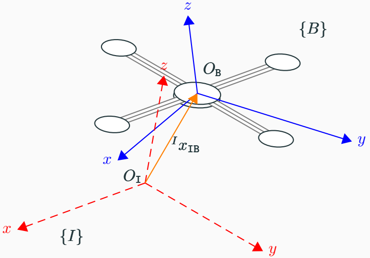
~~Continuum robot~~

The Pose of a Rigid Body[1]

Pose: "relative difference" between two frames, including

- Position: ${}^{I}x_{\mathtt{IB}} \in \mathbb{R}^3$
- Orientation: ${}^{I}R_B \in \mathrm{SO}(3)$ has three degrees of freedom

- 3D Rotation Group: orthonormal 3×3 matrices, i.e. ${}^{I}R_B {}^{I}R_B^\top = I$ with $\det = +1$
- Obtained from three **principle** rotations (those about one of the coordinate axes) with the Euler angles θ, ψ, γ:



$$
{}^{I}R_B(\theta, \psi, \gamma) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & \sin\theta \\ 0 & -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\gamma & \sin\gamma & 0 \\ -\sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}
$$

[1] $\{I\}$: Inertial frame, $\{B\}$: Body-fixed frame

Transformation Matrix

With the pose $({}^{I}x_{\mathtt{IB}}, {}^{I}R_{B})$, the coordinate p of a point A can be transformed from one frame to another:

$$ {}^{I}p = {}^{I}x_{\mathtt{IB}} + {}^{I}R_{B}\,{}^{B}p \quad\Longleftrightarrow\quad {}^{B}p = {}^{B}x_{\mathtt{BI}} + {}^{B}R_{I}\,{}^{I}p = -{}^{I}R_{B}^{\top}\,{}^{I}x_{\mathtt{IB}} + {}^{I}R_{B}^{\top}\,{}^{I}p $$

or compactly, using the (4×4) transformation matrix:

$$ \begin{bmatrix} {}^{I}p \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} {}^{I}R_{B} & {}^{I}x_{\mathtt{IB}} \\ 0_{1\times3} & 1 \end{bmatrix}}_{{}^{I}T_{B}} \begin{bmatrix} {}^{B}p \\ 1 \end{bmatrix}, \qquad {}^{I}T_{B} \in SE(3). $$

- Special structure with 16 parameters: Only 6 degrees of freedom + 10 constraints (6 constraints from $T^{\top}T = 1$; the other from the bottom row)
- Compound transformation matrices (it does not commute!)

$$ {}^{A}T_{C} = {}^{A}T_{B}\,{}^{B}T_{C} $$

Rigid-body Kinematics

The linear kinematics of ${}^I x := {}^I x_{\mathtt{IB}}$ can be obtained from basic geometry:

$$ {}^I\dot{x} = {}^I v = {}^I R_B {}^B v. \tag{1} $$

By studying the infinitesimal rotation, we can get the rotational kinematics of $R := {}^I R_B$ (Poisson's equation):

$$ \dot{R} = R\omega_\times, \qquad \omega_\times := \begin{bmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{bmatrix} \tag{2} $$

where the both the rotational velocity $\omega \in \mathbb{R}^3$ and the linear velocity $v \in \mathbb{R}^3$ are measured in the body frame $\{B\}$.

Pose kinematics (1)-(2) can be compactly written as

$$ {}^I\dot{T}_B = {}^I T_B \begin{bmatrix} \omega_\times & v \\ 0_{3\times 1} & 0 \end{bmatrix}. \tag{3} $$

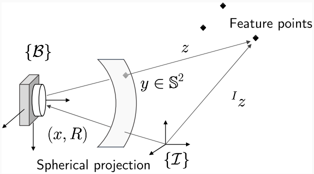
Robot Estimation: A Motivating Example

Mobile robots satisfy the dynamical model:

$$ {}^I\dot{T}_B = {}^I T_B \begin{bmatrix} \omega_\times & v \\ 0_{3\times 1} & 0 \end{bmatrix}, \quad {}^I T_B = \begin{bmatrix} R & x \\ 0_{3\times 1} & 1 \end{bmatrix} $$

with the velocities $\omega, v \in \mathbb{R}^3$ measured by the IMU in $\{B\}$.

- The internal state ${}^I T_B$ is unknown.
- The position of a feature point ${}^I z$ is unknown.



Output: A single monocular camera provides the bearing (direction) of the feature point in $\{B\}$:

$$ y = \frac{z}{|z|} \in \mathbb{S}^2, \qquad z := {}^B z = R^\top({}^I z - x). $$

Robotic Estimation from Monocular Camera. Given the input (ω, v), the output y, and the above dynamical model, how can we estimate the depth $|z(t)|$, or the robotic pose ${}^I T_B(t)$ in real-time?

The State Estimation Problem

The State Estimation Problem of a Probabilistic Robot

The dynamics and measurement of a robot are described by

$$\dot{x}(t) = f(x(t), u(t)) + w(t) \tag{4}$$

$$y(t) = g(x(t)) + v(t) \tag{5}$$

where at time t, $x(t) \in \mathbb{R}^n$ is the robot state, $u(t) \in \mathbb{R}^m$ is the control input, $y(t) \in \mathbb{R}^d$ is the sensor measurement, $w(t) \in \mathbb{R}^n$ is the disturbance, and $v(t) \in \mathbb{R}^d$ is the measurement noise.

Suppose at time t we know

- **Physics**: The system model (4) and (5).
- **Data**: The measured output $y(t)$ and control inputs $u(t)$ for $[0, t]$.
- **Uncertainty Characteristics**: Statistical distributions or bounds on $w(t), v(t)$.

The Problem: What is the most likely state $x(t)$ of the system?

State Estimator for Linear Systems: Observers

For linear systems in the state-space form of

$$
\begin{aligned}
\dot{x}(t) &= Ax(t) + Bu(t) + w(t) \\
y(t) &= Cx(t) + v(t)
\end{aligned}
$$

a state estimator can be constructed as the following observer form:

$$
\frac{d}{dt}\hat{x}(t) = \underbrace{A\hat{x}(t) + Bu(t)}_{\text{Model prediction}} + \underbrace{L(y(t) - C\hat{x}(t))}_{\text{Measurement correction}}
$$

The idea is that the gain L is chosen to balance between:

1. predictions based on the model and the current estimate $\hat{x}(t)$
2. new information arriving in the measurement $y(t)$.

This $\hat{x}(t)$ attempts to get close to $x(t)$!

Example: Velocity Observer

Consider the point-mass system $m\ddot{q} = 0$ again with $y = q$ as the system output. Set $x = [q \quad \dot{q}]^T$ then

$$\dot{x} = Ax = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x, \qquad y = Cx = \begin{bmatrix} 1 & 0 \end{bmatrix} x.$$

Let the observer gain matrix

$$L = \begin{bmatrix} l_1 \\ l_2. \end{bmatrix}$$

The observer is

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + L(y(t) - C\hat{x}(t)) && (6) \\ &= \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \hat{x}(t) + \begin{bmatrix} l_1 \\ l_2 \end{bmatrix} \left(y(t) - \begin{bmatrix} 1 & 0 \end{bmatrix} \hat{x}(t)\right) && (7) \end{aligned}$$

Estimation Error

Let us first consider the case where there is no measurement errors or disturbances. We define $\tilde{x} = x - \hat{x}$.

- We now have two differential equations

$$\dot{x}(t) = Ax(t) + Bu(t) + w(t) \tag{8}$$

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + L(y(t) - C\hat{x}(t)) \tag{9}$$

Noting $y(t) = Cx(t) + v(t)$, we have

$$\begin{aligned}
\frac{d\tilde{x}}{dt} &= \mathrm{RHS}(5) - \mathrm{RHS}(6) \\
&= Ax + Bu + w - (A\hat{x} + Bu + L(Cx + v - C\hat{x})) \\
&= A(x - \hat{x}) + \cancel{Bu} - \cancel{Bu} - LC(x - \hat{x}) + w + Lv \\
&= (A - LC)\tilde{x} + (w + Lv)
\end{aligned}$$

If eigenvalues of $A - LC$ have negative real parts, $\tilde{x}(t)$ will converge to some steady-state distributions.

Problem Redefinition: Discretization

A discrete-time probabilistic robot is described by

$$x_{k+1} = f(x_k, u_k) + w_k \tag{10}$$

$$y_k = g(x_k) + v_k \tag{11}$$

Suppose at step k we know

- **Physics**: The system model (10) and (11).
- **Data**: The measured output $y(s), s = 0, 1, \ldots, k$ and control inputs $u(s), s = 0, 1, \ldots, k-1$.
- **Uncertainty Characteristics**: Statistical distributions or bounds on each $w(s), v(s)$.

The Problem: What is our best estimation $\hat{x}_k$ for x_k?

Problem Redefinition: Linearization

A linearized robot is described by

$$x_{k+1} = A_k x_k + B_k u_k + w_k \tag{12}$$

$$y_k = C_k x_k + v_k. \tag{13}$$

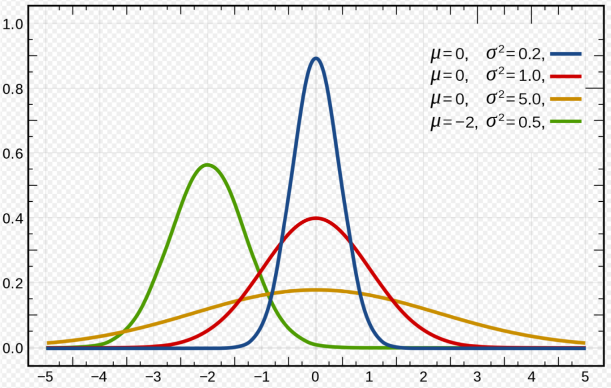
Now we are going to solve this linear state estimation problem!

The Kalman Filter

Gaussian Distributions

$$\mathcal{N}(\mu, \sigma^2) \sim \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{(x-\mu)^2}{2\sigma^2}}$$

- μ is the expectation (mean) capturing average from many independent trials.
- σ^2 is the variance capturing dispersion from the mean.



Linear Regressions

$$y = \mathbf{A}\beta$$

Linear Regressions: Minmum MSE Estimator

Let us consider following linear model

$$y = \mathbf{A}\beta + \epsilon \tag{14}$$

where $y \in \mathbb{R}^\ell$ is our measurement, $\beta \in \mathbb{R}^h$ is unknown parameter, and $\epsilon \in \mathbb{R}^\ell$ is random noise.

Assumption. $\mathbb{E}(\epsilon) = 0$ and $\mathbb{V}(\epsilon) = \mathbb{E}(\epsilon\epsilon^\top) = \Sigma \in \mathbb{R}^{\ell\times\ell}$.

An unbiased linear estimator $\hat{\beta}$ for β takes the form $\hat{\beta} = \mathbf{K}y$ so that $\mathbb{E}[\hat{\beta}] = \beta$. The resulting mean-square error is

$$\mathbb{MSE} = \mathbb{E}\big[(\hat{\beta} - \beta)(\hat{\beta} - \beta)^\top\big] = \mathbb{V}(\hat{\beta}) := \mathbf{K}\Sigma\mathbf{K}^\top.$$

Gauss-Markov Theorem. The linear, unbiased estimator that achieves the minimum mean-square error is given by

$$\hat{\beta}^* = (\mathbf{A}^\top\Sigma^{-1}\mathbf{A})^{-1}\mathbf{A}\Sigma^{-1}y$$

Linear Regressions: Maximum Likelihood Estimator

Let us again consider the model

$$y = \mathbf{A}\beta + \epsilon \tag{15}$$

under the following Gaussian assumption.

Assumption. $\epsilon \sim \mathcal{N}(0, \Sigma)$.

The likelihood of the measurement y under (deterministic) β can be computed as

$$p(y; \beta) = \frac{1}{\sqrt{(2\pi)^{\ell} \det(\Sigma)}} \exp\Big(-\frac{1}{2}(y - \mathbf{A}\beta)^{\top}\Sigma^{-1}(y - \mathbf{A}\beta)\Big)$$

The maximum likelihood estimator is given by

$$\hat{\beta}^{*} = \arg\max_{\beta}\ p(y; \beta) = (\mathbf{A}^{\top}\Sigma^{-1}\mathbf{A})^{-1}\mathbf{A}\Sigma^{-1}y$$

Linear Regressions: Bayesian Estimator

Now let us re-examine the model

$$y = \mathbf{A}\beta + \epsilon \tag{16}$$

and consider β as an unknown random variable independent with ϵ.

Assumption. The prior distributions satisfy $\epsilon \sim \mathcal{N}(0, \Sigma)$ and $\beta \sim \mathcal{N}(\beta_0, \sigma^2 I)$.

The posterior of β given y can be computed by Bayesian rule as

$$p(\beta|y) = \frac{p(y|\beta)p(\beta)}{p(y)} \sim \mathcal{N}(\widehat{\beta}, \widehat{V})$$

where

$$\widehat{\beta} = \beta_0 + V_{\beta y}V_y^{-1}(y - A\beta_0), \quad \widehat{V} = \sigma^2 I - V_{\beta y}V_y^{-1}V_{y\beta}$$

with $V_{\beta y} = \mathbb{E}(\beta y^\top) = \sigma^2 \mathbf{A}^\top, \quad V_y = \mathbb{E}(yy^\top) = \sigma^2 \mathbf{A}\mathbf{A}^\top + \Sigma$.

Linear Regressions: Bayesian Estimator

The Bayesian estimator under expected posterior quadratic loss is given by

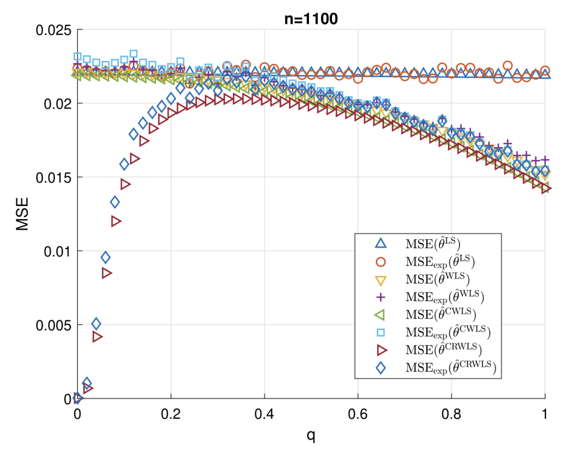
$$\begin{aligned}\hat{\beta}^* &= \arg\min_{\hat{\beta}} \int \|\hat{\beta} - \beta\|^2 p(\beta|y) d\beta \\ &= \int \beta p(\beta|y) \mathrm{d}\beta \\ &= \mathbb{E}(\beta|y) \\ &= \widehat{\beta} = \beta_0 + V_{\beta y} V_y^{-1} (y - A\beta_0)\end{aligned}$$

The MAP estimator is given by

$$\hat{\beta}^* = \arg\max_{\beta} p(\beta|y) = \widehat{\beta}.$$

Linear Regressions for Quantum State Estimation

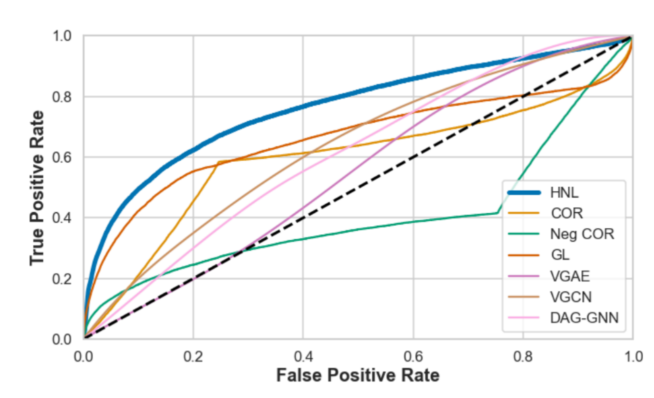
Werner State: $\rho = q|-\rangle\langle-| + \frac{1-q}{4} I$



[Mu, Qi, Petersen, Shi 2020]

Linear Regressions for Social Network Estimation

Learning hidden social influence from Los Angeles Yelp data.



[Leng, Chen, Dong, Wu, and Shi 2023]

Key Observations

Linear-Gaussian Model:

$$y = \mathbf{A}\beta + \epsilon$$

Assumption. (i) ϵ is independent with β; (ii) the prior distributions satisfy $\epsilon \sim \mathcal{N}(0, P_\epsilon)$ and $\beta \sim \mathcal{N}(\beta_0, P_\beta)$.

Observation 1. The conditional mean $\mathbb{E}(\beta|y)$ is the optimal Bayesian estimator for β upon observing y, that minimizes the expected posterior quadratic loss.

Observation 2. The $\mathbb{E}(\beta|y)$ and $\mathbb{V}(\beta|y)$, which carry all information about the Gaussian random variable $p(\beta|y)$, can be explicitly computed from $\beta_0, \mathbf{A}, P_\epsilon, P_\beta$.

The Optimal Filtering Problem

Now, let us come back to our robotic system and without loss of generality focus on

$$\begin{aligned} x_{k+1} &= A_k x_k + w_k \\ y_k &= C_k x_k + v_k. \end{aligned}$$

Assumption. (i) $\{v_k\}$ and $\{w_k\}$ are independent, zero mean, Gaussian with $\mathbb{E}(v_k v_k^\top) = R_k$ and $\mathbb{E}(w_k w_k^\top) = Q_k$;

(ii) The initial state x_0 is a Gaussian random variable with mean $\tilde{x}_0$ and covariance P_0, independent of $\{v_k\}$ and $\{w_k\}$.

Denoting $Y_k = \{y_0, \ldots, y_k\}$, the estimates

$$\hat{x}_{k|k-1} = \mathbb{E}(x_k | Y_{k-1}), \quad \hat{x}_{k|k} = \mathbb{E}(x_k | Y_k)$$

are the optimal prediction of x_k at step $k-1$, and the optimal estimation of x_k at step k in the Bayesian sense.

The Recursion: Step 1

Let us start from the initial time 0. What we have is

$$y_0 = C_0 x_0 + v_0.$$

This is exactly our standard Linear-Gaussian model, which leads to

$$\begin{aligned}\hat{x}_{0|0} &= \tilde{x}_0 + P_0 C_0^\top (C_0 P_0 C_0^\top + R_0)^{-1} (y_0 - C_0 \tilde{x}_0) \\ \Sigma_{0|0} &= P_0 - P_0 C_0^\top (C_0 P_0 C_0^\top + R_0)^{-1} C_0 P_0\end{aligned}$$

for the optimal Bayesian estimator and the corresponding covariance.

The Recursion: Step 2

Now at $k = 1$, from the robot dynamics

$$x_1 = A_0 x_0 + w_0 \tag{17}$$

This is again our standard Linear-Gaussian model. Direct computation gives us

$$\begin{aligned} \hat{x}_{1|0} &= A_0 \hat{x}_{0|0} \\ \Sigma_{1|0} &= A_0 \Sigma_{0|0} A_0^\top + Q_0. \end{aligned}$$

Before we receive y_1, $p(x_1|y_0)$ is our prior for x_1 which is Gaussian with mean $\hat{x}_{1|0}$ and variance $\Sigma_{1|0}$.

The Recursion: Step 3

At $k = 1$, from the measurement we have

$$y_1 = C_1 x_1 + v_1 \tag{18}$$

with $p(x_1|y_0) \sim \mathcal{N}(\hat{x}_{1|0}, \Sigma_{1|0})$ and $p(v_1) \sim \mathcal{N}(0, R_1)$.

For the third time, this is our standard Linear-Gaussian model, which gives us

$$\begin{aligned}
\hat{x}_{1|1} &= \hat{x}_{1|0} + \Sigma_{1|0} C_1^\top (C_1 \Sigma_{1|0} C_1^\top + R_1)^{-1} (y_1 - C_1 \hat{x}_{1|0}) \\
\Sigma_{1|1} &= \Sigma_{1|0} - \Sigma_{1|0} C_1^\top (C_1 \Sigma_{1|0} C_1^\top + R_1)^{-1} C_1 \Sigma_{1|0}.
\end{aligned}$$

We can now move to $\hat{x}_{2|1}$ and $\Sigma_{2|1}$ and the computation starts to repeat itself recursively!

The Kalman Filter

The system model

$$x_{k+1} = A_k x_k + B_k u_k + w_k$$
$$y_k = C_k x_k + v_k$$

under Gaussian noises admits the following optimal filtering solution, namely the Kalman filter:

$$\begin{aligned}
\text{State prediction}: &\quad \hat{x}_{k|k-1} = A_{k-1} x_{k-1|k-1} + B_{k-1} u_{k-1} \\
\text{Prediction covariance}: &\quad \Sigma_{k|k-1} = A_{k-1} \Sigma_{k-1|k-1} A_{k-1}^\top + Q_{k-1} \\
\text{Innovation}: &\quad \tilde{z}_k = y_k - C_k \hat{x}_{k|k-1} \\
\text{Innovation Covariance}: &\quad S_k = C_k \Sigma_{k|k-1} C_k^\top + R_k \\
\text{Kalman Gain}: &\quad K_k = \Sigma_{k|k-1} C_k^\top S_k^{-1} \\
\text{State Estimate}: &\quad \hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{z}_k \\
\text{Estimate covariance}: &\quad \Sigma_{k|k} = (I - K_k C_{k-1}) \Sigma_{k|k-1}
\end{aligned}$$

The Extended Kalman Filter

For the nonlinear probabilistic robot system

$$x_{k+1} = f(x_k, u_k) + w_k$$
$$y_k = g(x_k) + v_k$$

suppose we have obtained $\hat{x}_{k-1|k-1}$ and $\Sigma_{k-1|k-1}$.

- The state prediction can be estimated as $\hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1}, u_{k-1})$.
- To estimate the prediction covariance, we can use

$$\begin{aligned} x_k &\approx f(\hat{x}_{k-1|k-1}, u_{k-1}) + A_{k-1}(x_{k-1} - \hat{x}_{k-1|k-1}) + w_{k-1} \\ &= \hat{x}_{k-1|k-1} + A_{k-1}(x_{k-1} - \hat{x}_{k-1|k-1}) + w_{k-1} \end{aligned}$$

where $A_{k-1} = \frac{\partial}{\partial x} f(x, u)\Big|_{(\hat{x}_{k-1|k-1}, u_{k-1})}$.

The Extended Kalman Filter

For the nonlinear probabilistic robot system

$$x_{k+1} = f(x_k, u_k) + w_k$$
$$y_k = g(x_k) + v_k$$

the extended Kalman filter is given by

$$\begin{aligned}
\text{State prediction}: &\quad \hat{x}_{k|k-1} = f(\hat{x}_{k-1|k-1}, u_{k-1}) \\
\text{Prediction covariance}: &\quad \Sigma_{k|k-1} = A_{k-1}\Sigma_{k-1|k-1}A_{k-1}^\top + Q_{k-1} \\
\text{Innovation}: &\quad \tilde{z}_k = y_k - h(\hat{x}_{k|k-1}) \\
\text{Innovation Covariance}: &\quad S_k = C_k\Sigma_{k|k-1}C_k^\top + R_k \\
\text{Kalman Gain}: &\quad K_k = \Sigma_{k|k-1}C_k^\top S_k^{-1} \\
\text{State Estimate}: &\quad \hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k\tilde{z}_k \\
\text{Estimate covariance}: &\quad \Sigma_{k|k} = (I - K_kC_{k-1})\Sigma_{k|k-1}
\end{aligned}$$

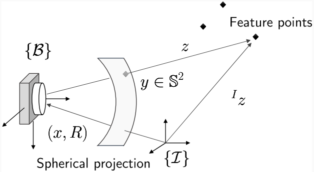
The Motivating Example Revisited

Mobile robots satisfy the dynamical model:

$$ {}^{I}\dot{T}_B = {}^{I}T_B \begin{bmatrix} \omega_\times & v \\ 0_{3\times 1} & 0 \end{bmatrix}, \quad {}^{I}T_B = \begin{bmatrix} R & x \\ 0_{3\times 1} & 1 \end{bmatrix} $$

with the velocities $\omega, v \in \mathbb{R}^3$ measured by the IMU in $\{B\}$. A single monocular camera provides the bearing (direction) of the feature point in $\{B\}$:

$$ y = \frac{z}{|z|} \in \mathbb{S}^2, \qquad z := {}^{B}z = R^\top({}^{I}z - x). $$



- EKF is in place to give us a solution where we only need to augment the dynamics of z into the system equations.
  - Estimating ${}^{I}z$ is a parameter estimation problem.
  - Estimating ${}^{I}T_B$ is a localization (navigation) problem.
  - Estimating both z and ${}^{I}T_B$ is a SLAM problem.
- By exploiting the geometry of the problem, nonlinear observers solve the same problem with global convergence guarantee for z [Yi, Jin, and Manchester 2022].

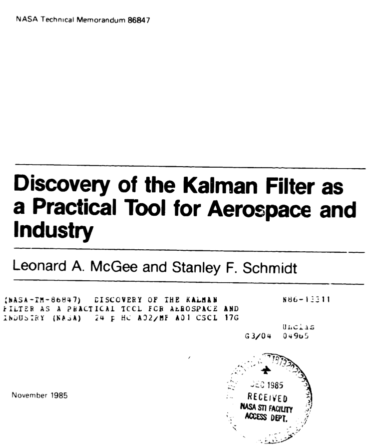
Conclusions

Summary

- Robotic estimation is to estimate unknown quantities by combining the physics, data, and uncertainty characteristics that we know about the robot.
- Robotic filtering is to estimate the robotic state in real-time from the information available at each moment.
- Kalman filter provides a recursive algorithm for robotic filtering that is optimal in the Bayesian sense for linear-Gaussian settings; extended Kalman filter allows us to continue to use the algorithm for nonlinear systems.
- The problem of estimating a future state based on currently available information is called prediction; the problem of estimating past states with current information is called smoothing.

Historical Remarks

- The Kalman filter framework was published in 1960 from a seminal paper by Rudolf E. Kalman.
- Kalman filter was used on the Apollo 11 lunar module for the estimation of the module's position above the lunar surface based on noisy radar measurements.



NASA Technical Memorandum 86847

Discovery of the Kalman Filter as a Practical Tool for Aerospace and Industry

Leonard A. McGee and Stanley F. Schmidt

(NASA-TM-86847) DISCOVERY OF THE KALMAN FILTER AS A PRACTICAL TOOL FOR AEROSPACE AND INDUSTRY (NASA) 24 p HC A02/MF A01 CSCL 17G

N86-13311

Unclas

G3/04 04965

DEC 1985
RECEIVED
NASA STI FACILITY
ACCESS DEPT.

November 1985

Brian Anderson - 1992 recipient of IEEE CSS Hendrik W. Bode Lecture Prize

Award Received:

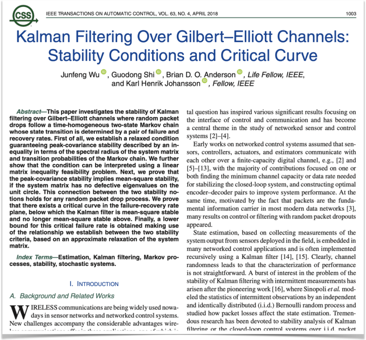
IEEE CSS Hendrik W. Bode Lecture Prize

Date Awarded: 1992

Brian Anderson

Citation:

For revealing fundamental feedback properties in scholarly papers on optimal and adaptive filtering and for seminal expository textbooks and monographs



IEEE TRANSACTIONS ON AUTOMATIC CONTROL, VOL. 63, NO. 4, APRIL 2018 1003

Kalman Filtering Over Gilbert–Elliott Channels: Stability Conditions and Critical Curve

Junfeng Wu, Guodong Shi, Brian D. O. Anderson, *Life Fellow, IEEE*, and Karl Henrik Johansson, *Fellow, IEEE*

***Abstract*—This paper investigates the stability of Kalman filtering over Gilbert–Elliott channels where random packet drops follow a time-homogeneous two-state Markov chain whose state transition is determined by a pair of failure and recovery rates. First of all, we establish a relaxed condition guaranteeing peak-covariance stability described by an inequality in terms of the spectral radius of the system matrix and transition probabilities of the Markov chain. We further show that the condition can be interpreted using a linear matrix inequality feasibility problem. Next, we prove that the peak-covariance stability implies mean-square stability, if the system matrix has no defective eigenvalues on the unit circle. This connection between the two stability notions holds for any random packet drop process. We prove that there exists a critical curve in the failure-recovery rate plane, below which the Kalman filter is mean-square stable and no longer mean-square stable above. Finally, a lower bound for this critical failure rate is obtained making use of the relationship we establish between the two stability criteria, based on an approximate relaxation of the system matrix.**

***Index Terms*—Estimation, Kalman filtering, Markov processes, stability, stochastic systems.**

I. Introduction

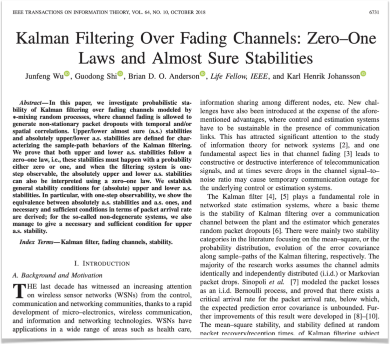
A. Background and Related Works

WIRELESS communications are being widely used nowadays in sensor networks and networked control systems. New challenges accompany the considerable advantages wire-

tal question has inspired various significant results focusing on the interface of control and communication and has become a central theme in the study of networked sensor and control systems [2]–[4].

Early works on networked control systems assumed that sensors, controllers, actuators, and estimators communicate with each other over a finite-capacity digital channel, e.g., [2] and [5]–[13], with the majority of contributions focused on one or both finding the minimum channel capacity or data rate needed for stabilizing the closed-loop system, and constructing optimal encoder–decoder pairs to improve system performance. At the same time, motivated by the fact that packets are the fundamental information carrier in most modern data networks [3], many results on control or filtering with random packet dropouts appeared.

State estimation, based on collecting measurements of the system output from sensors deployed in the field, is embedded in many networked control applications and is often implemented recursively using a Kalman filter [14], [15]. Clearly, channel randomness leads to that the characterization of performance is not straightforward. A burst of interest in the problem of the stability of Kalman filtering with intermittent measurements has arisen after the pioneering work [16], where Sinopoli *et al.* modeled the statistics of intermittent observations by an independent and identically distributed (i.i.d.) Bernoulli random process and studied how packet losses affect the state estimation. Tremendous research has been devoted to stability analysis of Kalman filtering or the closed-loop control systems over i.i.d. packet



IEEE TRANSACTIONS ON INFORMATION THEORY, VOL. 64, NO. 10, OCTOBER 2018 6731

Kalman Filtering Over Fading Channels: Zero–One Laws and Almost Sure Stabilities

Junfeng Wu, Guodong Shi, Brian D. O. Anderson, *Life Fellow, IEEE*, and Karl Henrik Johansson

***Abstract*—In this paper, we investigate probabilistic stability of Kalman filtering over fading channels modeled by ∗-mixing random processes, where channel fading is allowed to generate non-stationary packet dropouts with temporal and/or spatial correlations. Upper/lower almost sure (a.s.) stabilities and absolutely upper/lower a.s. stabilities are defined for characterizing the sample-path behaviors of the Kalman filtering. We prove that both upper and lower a.s. stabilities follow a zero–one law, i.e., these stabilities must happen with a probability either zero or one, and when the filtering system is one-step observable, the absolutely upper and lower a.s. stabilities can also be interpreted using a zero–one law. We establish general stability conditions for (absolute) upper and lower a.s. stabilities. In particular, with one-step observability, we show the equivalence between absolutely a.s. stabilities and a.s. ones, and necessary and sufficient conditions in terms of packet arrival rate are derived; for the so-called non-degenerate systems, we also manage to give a necessary and sufficient condition for upper a.s. stability.**

***Index Terms*—Kalman filter, fading channels, stability.**

I. Introduction

A. Background and Motivation

THE last decade has witnessed an increasing attention on wireless sensor networks (WSNs) from the control, communication and networking communities, thanks to a rapid development of micro–electronics, wireless communication, and information and networking technologies. WSNs have applications in a wide range of areas such as health care, information sharing among different nodes, etc. New challenges have also been introduced at the expense of the aforementioned advantages, where control and estimation systems have to be sustainable in the presence of communication links. This has attracted significant attention to the study of information theory for network systems [2], and one fundamental aspect lies in that channel fading [3] leads to constructive or destructive interference of telecommunication signals, and at times severe drops in the channel signal–to–noise ratio may cause temporary communication outage for the underlying control or estimation systems.

The Kalman filter [4], [5] plays a fundamental role in networked state estimation systems, where a basic theme is the stability of Kalman filtering over a communication channel between the plant and the estimator which generates random packet dropouts [6]. There were mainly two stability categories in the literature focusing on the mean–square, or the probability distribution, evolution of the error covariance along sample–paths of the Kalman filtering, respectively. The majority of the research works assumes the channel admits identically and independently distributed (i.i.d.) or Markovian packet drops. Sinopoli *et al.* [7] modeled the packet losses as an i.i.d. Bernoulli process, and proved that there exists a critical arrival rate for the packet arrival rate, below which, the expected prediction error covariance is unbounded. Further improvements of this result were developed in [8]–[10]. The mean–square stability, and stability defined at random packet recovery/reception times, of Kalman filtering subject