02427 Advanced Time Series Analysis

Lecture 12

Tobias K. S. Ritschel

Department of Applied Mathematics and Computer Science,
Technical University of Denmark

November 24, 2026

Outline

Stochastic control

Mixture approximations

Stochastic differential equations

Differential equations with distributed time delays

Stochastic control

Control problem

System

$$x_{t+1} = ax_t + bu_t + v_t, \qquad v_t \sim N(0, \sigma_v^2), \qquad x_0 \sim N(0, \sigma_0^2), \tag{1a}$$

$$y_t = x_t + e_t, \qquad e_t \sim N(0, \sigma_e^2) \tag{1b}$$

Optimal control problem

$$\min_{u_t} \quad \phi_t = \mathbb{E}[(y_{t+1} - \bar{y}_t)^2] \tag{2}$$

Future output

$$y_{t+1} = x_{t+1} + e_{t+1} = ax_t + bu_t + v_t + e_{t+1} \tag{3}$$

Expected future output and error

$$\begin{aligned} \hat{y}_{t+1|t} &= \mathbb{E}[y_{t+1}|y_t] = \mathbb{E}[ax_t + bu_t + v_t + e_{t+1}|y_t] \\ &= a\mathbb{E}[x_t|y_t] + bu_t = a\hat{x}_{t|t} + bu_t, \end{aligned} \tag{4a}$$

$$\tilde{y}_{t+1|t} = v_t + e_{t+1} \tag{4b}$$

Future output (prediction + error)

$$y_{t+1} = \hat{y}_{t+1|t} + \tilde{y}_{t+1|t} \tag{5}$$

Optimal control law

Rewrite objective function

$$\begin{aligned}\phi_t &= \mathbb{E}[(y_{t+1} - \bar{y}_t)^2] = \mathbb{E}[(\hat{y}_{t+1|t} + \tilde{y}_{t+1|t} - \bar{y}_t)^2] \\ &= \mathbb{E}[(\hat{y}_{t+1|t} - \bar{y}_t)^2 + \tilde{y}^2_{t+1|t} - 2(\hat{y}_{t+1|t} - \bar{y}_t)\tilde{y}_{t+1|t}] \\ &= \mathbb{E}[(\hat{y}_{t+1|t} - \bar{y}_t)^2] + \underbrace{\mathbb{E}[\tilde{y}^2_{t+1|t}]}_{=\sigma_v^2+\sigma_e^2} - 2(\hat{y}_{t+1|t} - \bar{y}_t)\underbrace{\mathbb{E}[\tilde{y}_{t+1|t}]}_{=0}\end{aligned} \tag{6}$$

Equivalent optimal control problem

$$\min_{u_t} \quad \hat{\phi}_t = (\hat{y}_{t+1|t} - \bar{y}_t)^2 \tag{7}$$

Solution: Differentiate and put equal to zero

$$\begin{aligned}\frac{\partial \hat{\phi}_t}{\partial u_t} &= 2(\hat{y}_{t+1|t} - \bar{y}_t)\frac{\partial \hat{y}_{t+1|t}}{\partial u_t} = 0 \\ &= 2(a\hat{x}_{t|t} + bu_t - \bar{y}_t)b\end{aligned} \tag{8}$$

Solution

$$u_t = \frac{\bar{y}_t - a\hat{x}_{t|t}}{b} \tag{9}$$

Written exercise

System

$$x_{t+1} = ax_t + bu_t + v_t, \qquad v_t \sim N(0, \sigma_v^2), \qquad x_0 \sim N(0, \sigma_0^2), \tag{10a}$$

$$y_t = x_t + e_t, \qquad e_t \sim N(0, \sigma_e^2) \tag{10b}$$

Optimal control law (e.g., $\hat{x}_{t|t} = y_t$, which is not a good estimate, but easy to implement for the sake of the example)

$$u_t = \frac{\bar{y}_t - \hat{a}\hat{x}_{t|t}}{b} \tag{11}$$

Parameter values

$$a = 0.95, \qquad b = 0.2, \qquad \bar{y}_t = 0.5, \tag{12a}$$

$$\sigma_v^2 = 10^{-4}, \qquad \sigma_e^2 = 10^{-4}, \qquad \sigma_0^2 = 0.1 \tag{12b}$$

Simulate $N = 10^3$ steps where

1. $\hat{a} = a = 0.95$
2. $\hat{a} = 0.9$ (plant-model mismatch)

What about some of the models from this course?

ARX system

$$y_t + a_1 y_{t-1} + \cdots a_n y_{t-n} = b_0 u_t + b_1 u_{t-1} + \cdots b_m u_{t-m} + e_t \tag{13}$$

Deterministic parameters

$$a_i = a_i(Y_{t-1}), \qquad Y_{t-1} = \left[y_{t-1}; y_{t-2}; \ldots\right], \tag{14a}$$
$$b_i = b_i(Y_{t-1}) \tag{14b}$$

or stochastic parameters

$$a_i = \mathcal{F}_i^a(Y_{t-1}), \tag{15a}$$
$$b_i = \mathcal{F}_i^b(Y_{t-1}) \tag{15b}$$

- How can we develop stochastic controllers for such systems?
- Can we also develop stochastic *adaptive* controllers for them?

Mixture approximations

Problem statement

Approximate the probability density function $p(x)$

1. Non-negative and bounded

$$0 \leq p(x) \leq K, \qquad x \in \mathcal{D}$$

2. Integrates to one

$$\int_{\mathcal{D}} p(x)\,\mathrm{d}x = 1$$

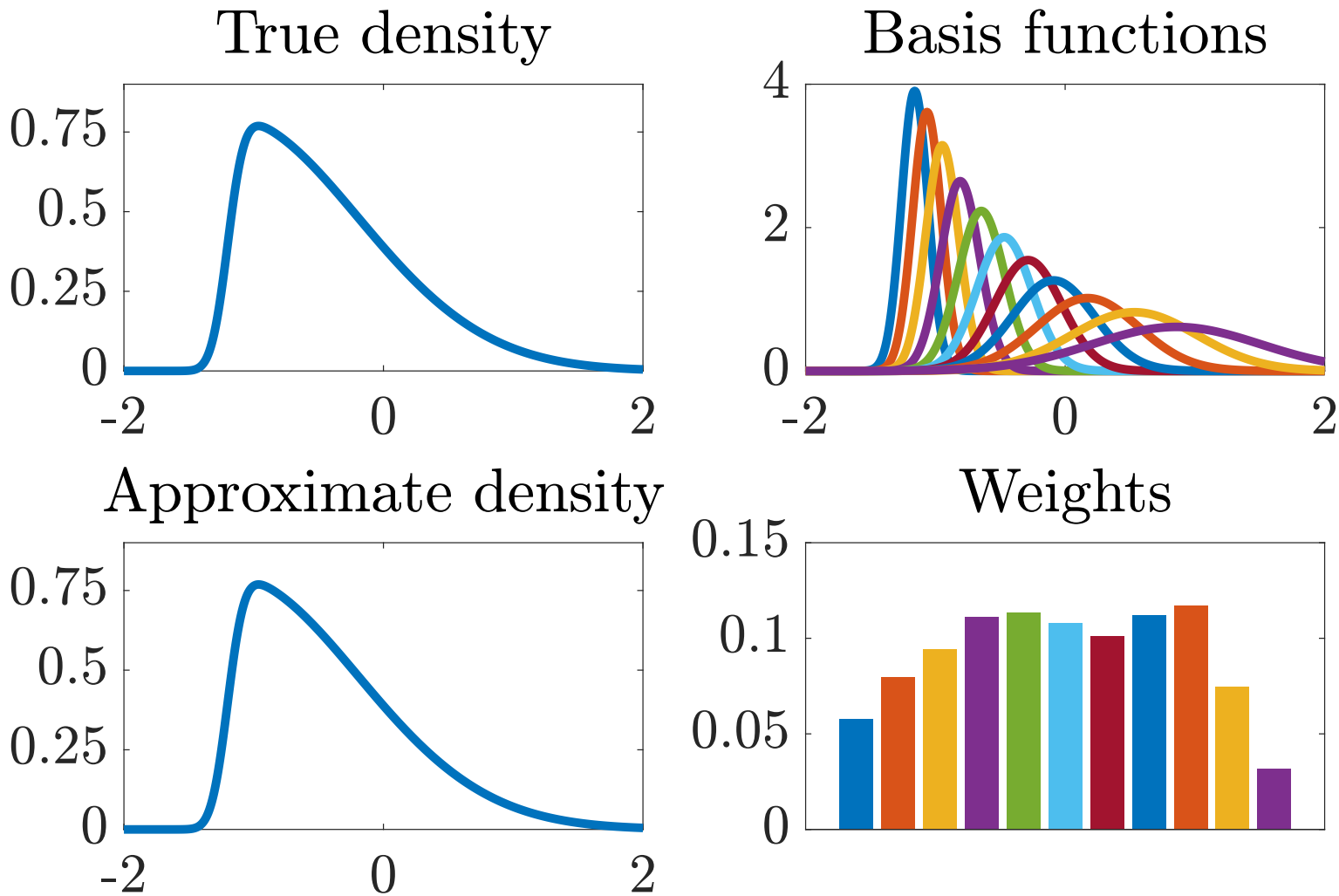
3. Continuous

The domain may be

- the real numbers, $\mathcal{D} = (-\infty, \infty)$,
- the non-negative real numbers, $\mathcal{D} = [0, \infty)$, or
- an interval, $\mathcal{D} = [x_0, x_f]$

Example for $\mathcal{D} = (-\infty, \infty)$

Skew normal distribution



Mixture approximation

Approximate p by a sum of basis functions

$$p(x) \approx \hat{p}(x) = \sum_{m=0}^{M} c_m \ell_m(x)$$

Weight constraints

$$\sum_{m=0}^{M} c_m = 1, \qquad 0 \leq c_m, \qquad m = 0, \ldots, M$$

Non-negative basis functions

$$\ell_m(x) \geq 0, \qquad m = 0, \ldots, M$$

Basis functions integrate to one

$$\int_{\mathcal{D}} \ell_m(x) \, \mathrm{d}x = 1, \qquad m = 0, \ldots, M$$

Gaussian, Erlang, and beta mixture approximations

Gaussian basis functions, $\mathcal{D} = (-\infty, \infty)$

$$\ell_m(x) = b_m \exp\left(-\frac{1}{2}\left(\frac{x-\mu_m}{\sigma_m}\right)^2\right), \qquad b_m = \frac{1}{\sqrt{2\pi\sigma_m^2}}$$

Erlang basis functions, $\mathcal{D} = [0, \infty)$

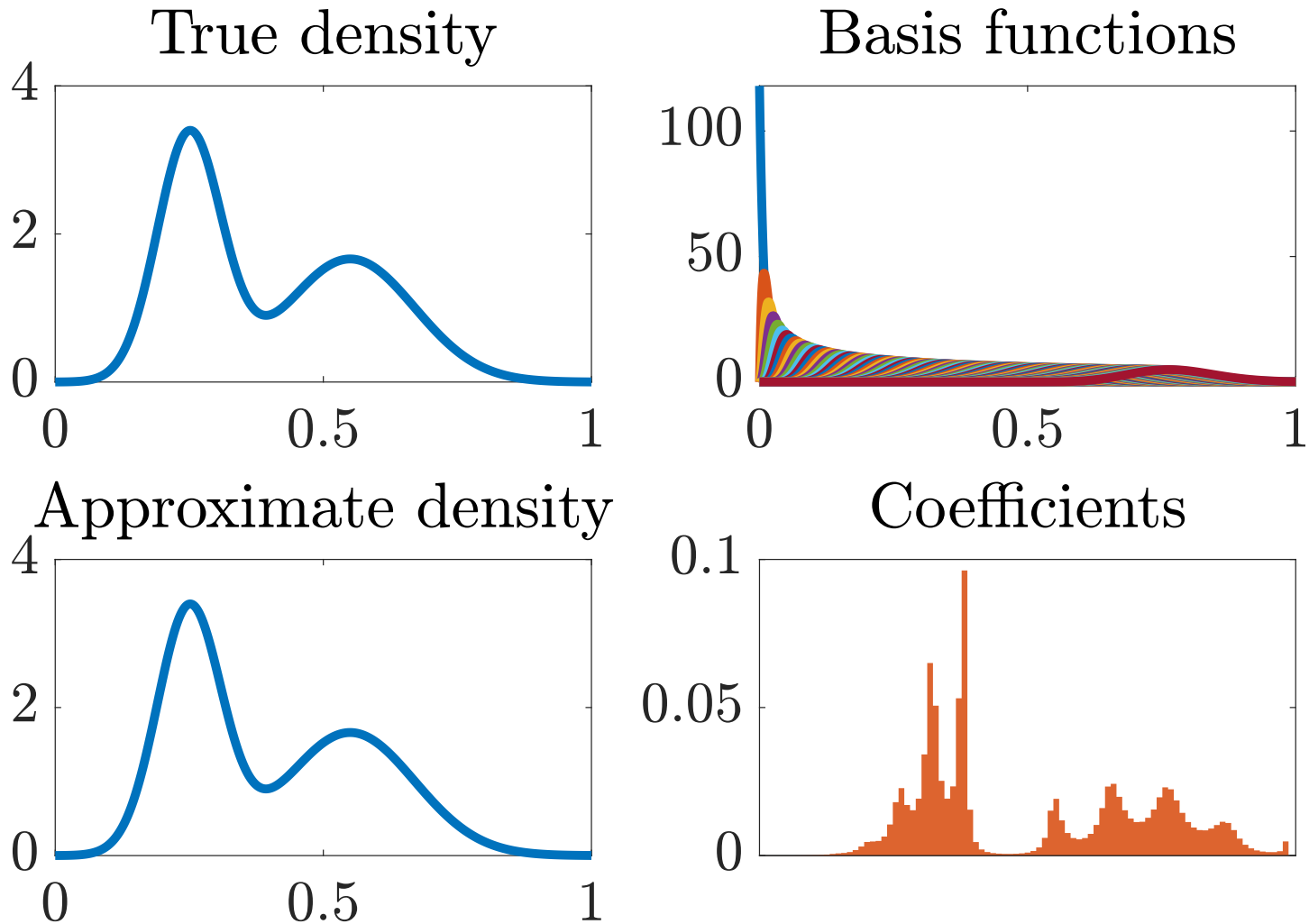
$$\ell_m(x) = b_m x^m e^{-ax}, \qquad b_m = \frac{a^{m+1}}{m!}$$

Beta basis functions, $\mathcal{D} = [0, x_f]$

$$\ell_m(x) = b_m x^m (x_f - x)^{M-m}, \qquad b_m = \frac{1}{x_f^{M+1}} \frac{(M+1)!}{m!(M-m)!}$$

Example: Erlang mixture approximation ($M = 90$)

Folded normal mixture distribution



Optimal coefficients and rate parameter ($\mathcal{D} = [0, \infty)$)

Domain ($k = 0, \ldots, N-1$)

$$\int_0^X p(x)\,\mathrm{d}x = 1 - \epsilon, \qquad\qquad x_k = k/X$$

Least-squares optimization problem

$$\begin{aligned} \min_{\{c_m\}_{m=0}^M, a} \quad & \phi = \frac{1}{2} \sum_{k=0}^{N-1} (p(x_k) - \hat{p}(x_k))^2, \\ \text{subject to} \quad & \sum_{m=0}^{M} c_m = 1, \\ & 0 \le c_m \le 1, \quad m = 0, \ldots, M, \\ & a_{\min} \le a \end{aligned}$$

Written exercise

Fit an Erlang mixture approximation to a histogram of the 9$^{\text{th}}$ column of `Breast_cancer_dataset.csv` with $N = 20$ bins and 'pdf' normalization.

Least-squares optimization problem (use bin centers as x_k)

$$\begin{aligned} \min_{\{c_m\}_{m=0}^{M}, a} \quad & \phi = \frac{1}{2} \sum_{k=0}^{N-1} (p(x_k) - \hat{p}(x_k))^2, \\ \text{subject to} \quad & \sum_{m=0}^{M} c_m = 1, \\ & 0 \le c_m \le 1, \quad m = 0, \ldots, M, \\ & a_{\min} \le a \end{aligned}$$

Erlang mixture approximation

$$\hat{p}(x) = \sum_{m=0}^{M} c_m \ell_m(x), \qquad \ell_m(x) = b_m x^m e^{-ax}, \qquad b_m = \frac{a^{m+1}}{m!} \tag{18}$$

Parameters (e.g., $M = 5$)

$$a_{\min} = 10^{-3}, \quad \text{Initial guess of } c_m = \frac{1}{M+1}, \quad \text{Initial guess of } a = 20$$

Stochastic differential equations

Stochastic differential equations

Stochastic differential equation

$$\mathrm{d}x(t) = f(x(t))\,\mathrm{d}t + g(x(t))\,\mathrm{d}w(t)$$

Solution (time-varying probability density function)

$$p(t, x) = ?$$

Fokker-Planck equation

$$\partial_t p(t, x) = -\partial_x(f(x)p(t, x)) + \frac{1}{2}\partial_{xx}(g^2(x)p(t, x))$$

Mixture approximation

$$p(t, x) \approx \hat{p}(t, x) = \sum_{m=0}^{M} c_m(t)\ell_m(x)$$

Substitute into Fokker-Planck equation

$$\begin{aligned} R(t, x) &= \sum_{m=0}^{M} \left(\ell_m(x)\dot{c}_m(t) + \left[\partial_x(f(x)\ell_m(x)) - \frac{1}{2}\partial_{xx}(g^2(x)\ell_m(x))\right] c_m(t) \right) \\ &= \sum_{m=0}^{M} \left(\ell_m(x)\dot{c}_m(t) + A_m(x)c_m(t) \right) \end{aligned}$$

Minimize residual function

Least-squares approach

$$\min_{\{\dot{c}_m\}_{m=0}^{M}} \quad \phi = \frac{1}{2} \int_{\mathcal{D}} R^2(t,x)\,\mathrm{d}x$$

First-order optimality conditions

$$\frac{\partial \phi}{\partial \dot{c}_j} = \sum_{m=0}^{M} \left(\int_{\mathcal{D}} \ell_m(x)\ell_j(x)\,\mathrm{d}x\ \dot{c}_m(t) + \int_{\mathcal{D}} \ell_j(x)A_m(x)\,\mathrm{d}x\ c_m(t) \right) = 0$$

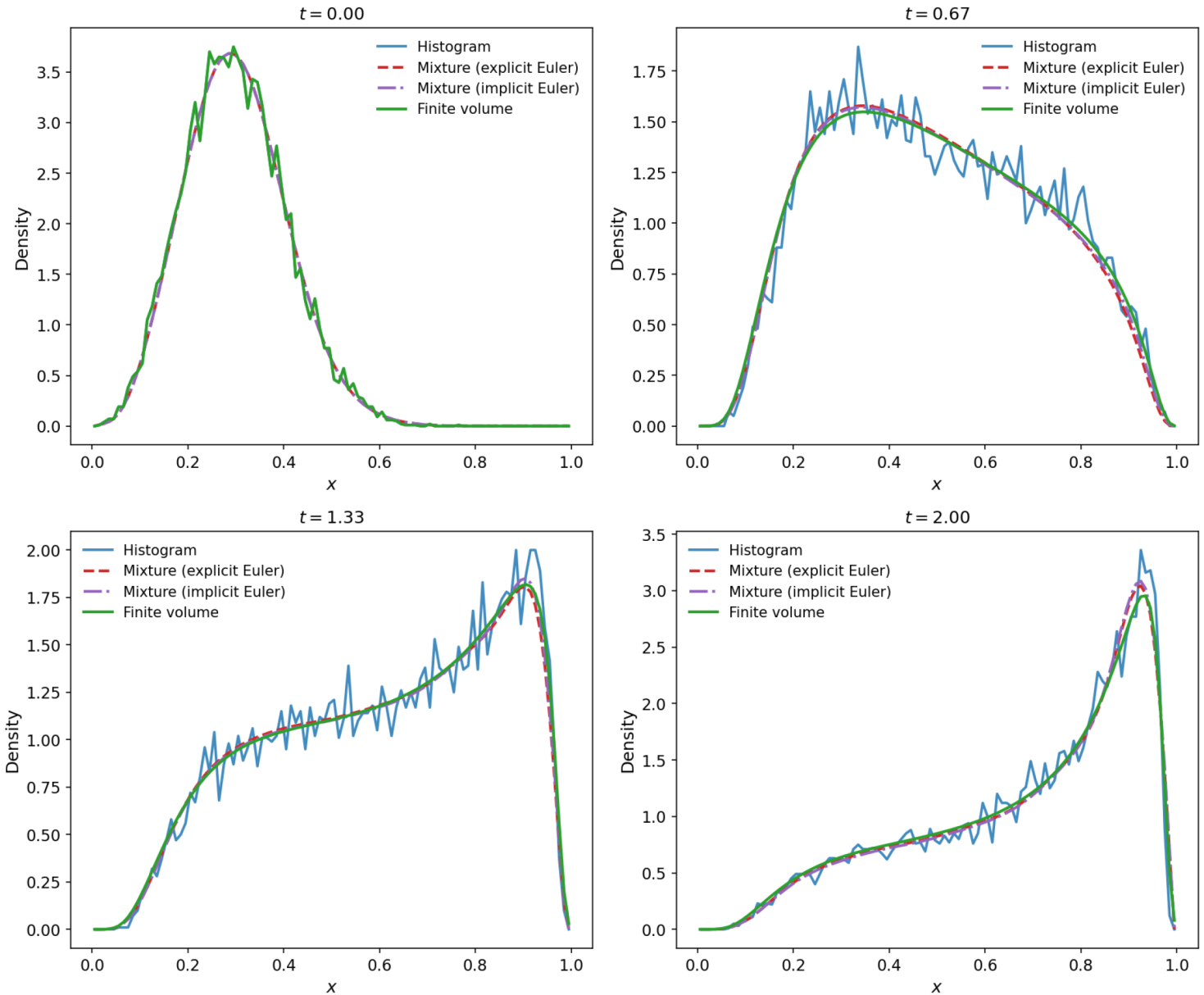
This is a linear system, and we can isolate the time derivatives

$$\dot{c}_j(t) = \sum_{m=0}^{M} w_{jm} c_m(t)$$

- How do we compute w_{jm}?
- How do we ensure that $0 \leq c_m$ and $\sum_{m=0}^{M} c_m(t) = 1$?
- How do we use it for state estimation?
- How do we use it for parameter estimation?
- How do we use it for stochastic optimal control?

Demand response model

MSc project by Alexander Elbæk Nielsen: Link to project.



Differential equations with distributed time delays

Delay differential equations

Logistic equation

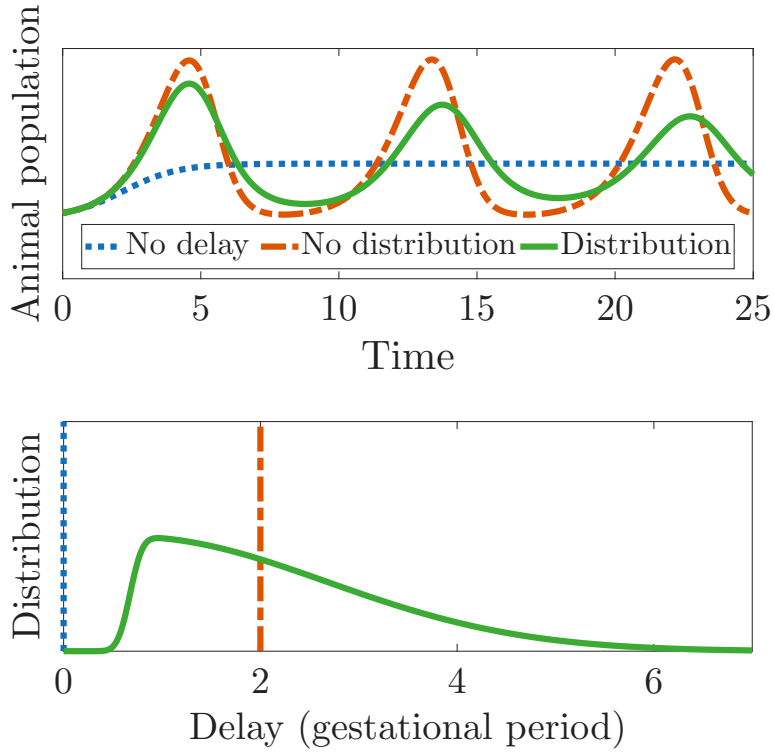
$$\dot{x}(t) = x(t)\,(1 - x(t))$$

Logistic equation (absolute time delay)

$$\dot{x}(t) = x(t)\,(1 - x(t - \tau))$$

Logistic equation (distributed time delay)

$$\dot{x}(t) = x(t)\left(1 - \int_{-\infty}^{t} \alpha(t - s)x(s)\,\mathrm{d}s\right)$$



Erlang mixture approximation of α

$$\alpha(t) \approx \hat{\alpha}(t) = \sum_{m=0}^{M} c_m \ell_m(t), \qquad \ell_m(t) = b_m t^m e^{-at} \tag{19}$$

- Can we estimate α (or the c_m's) from time series data?

Epidemiology

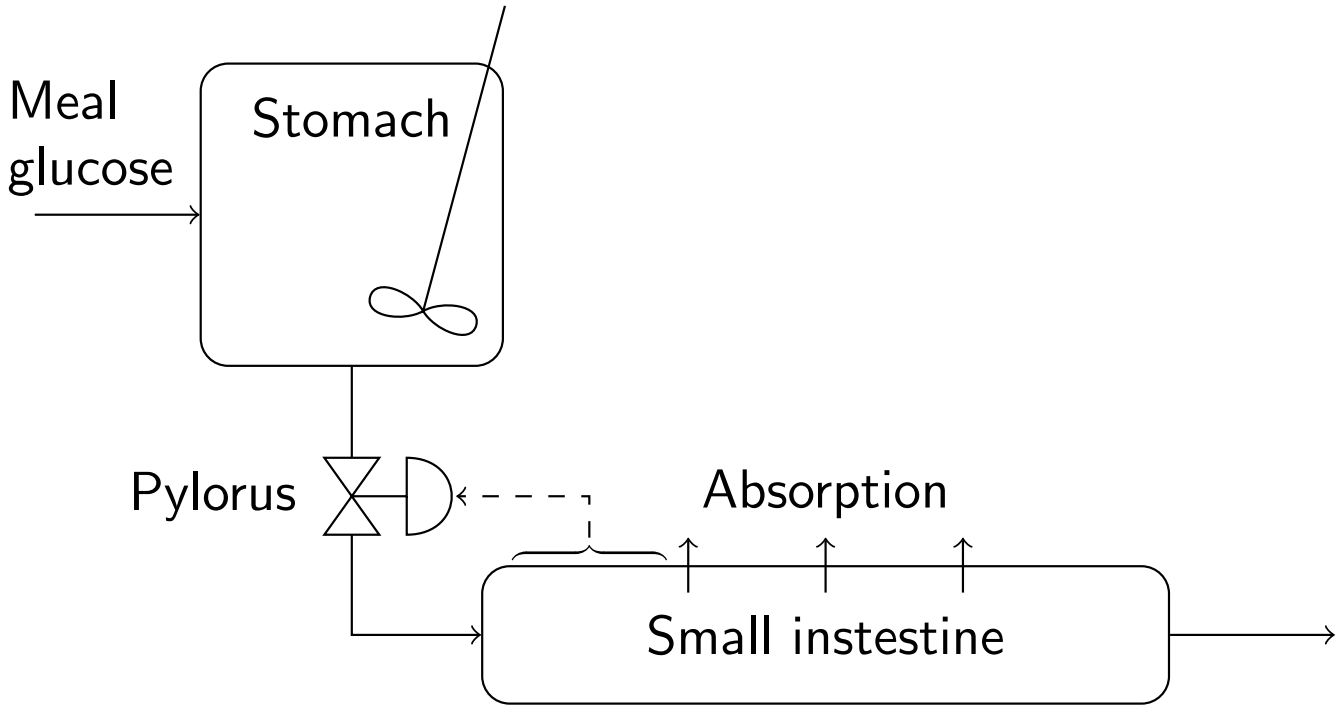
SIR model without delays

$$\begin{aligned}\dot{S}(t) &= -\beta S(t) I(t),\\ \dot{I}(t) &= \beta S(t) I(t) - \eta I(t),\\ \dot{R}(t) &= \eta I(t)\end{aligned}$$

SIR model with distributed time delay

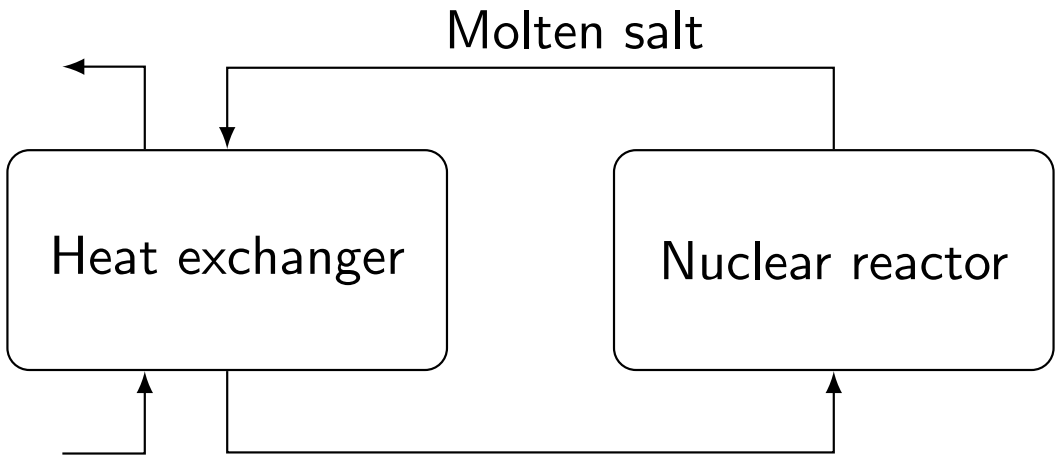
$$\begin{aligned}\dot{S}(t) &= -\beta S(t) \int_{-\infty}^{t} \alpha(t-s) I(s)\,\mathrm{d}s,\\ \dot{I}(t) &= \beta S(t) \int_{-\infty}^{t} \alpha(t-s) I(s)\,\mathrm{d}s - \eta I(t),\\ \dot{R}(t) &= \eta I(t)\end{aligned}$$

Diabetes

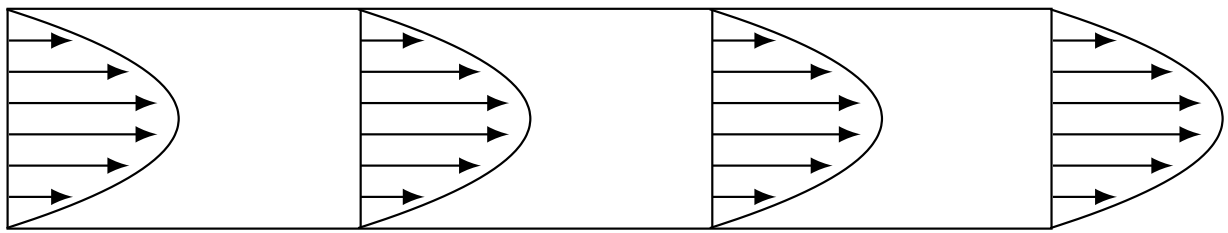


Ritschel, T.K.S., Reenberg, A.T., Carstensen, P.E., Bendsen, J., Jørgensen, J.B., 2023. Mathematical Meal Models for Simulation of Human Metabolism. arXiv: 2307.16444.

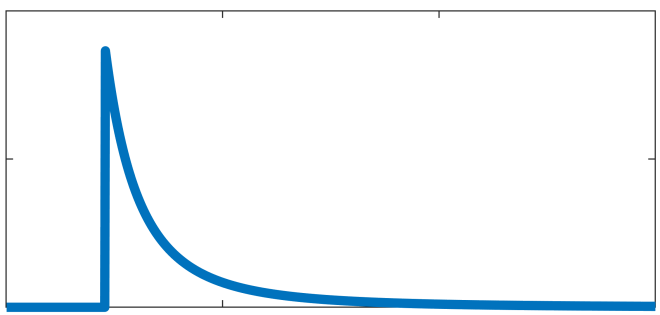
Molten salt nuclear reactor and nonuniform flow in pipes



Non-uniform velocity profile



Kernel for Hagen-Poiseuille flow (quadratic velocity profile)



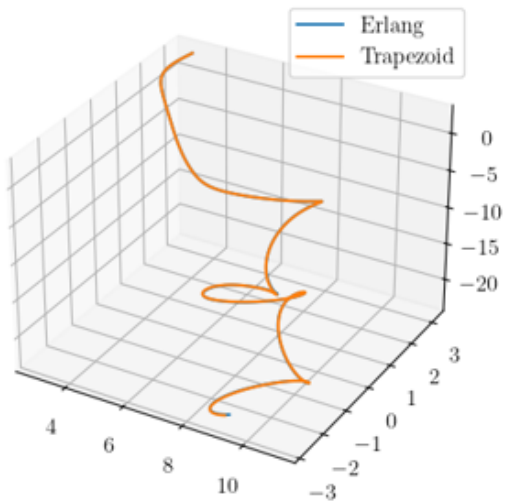
Ritschel, T.K.S., 2025. Numerical Optimal Control for Distributed Delay Differential Equations: A Simultaneous Approach based on Linearization of the Delayed Variables. In: Proceedings of the 2025 European Control Conference (ECC), June 24-27, Thessaloniki, Greece. DOI: 10.23919/ECC65951.2025.11187183

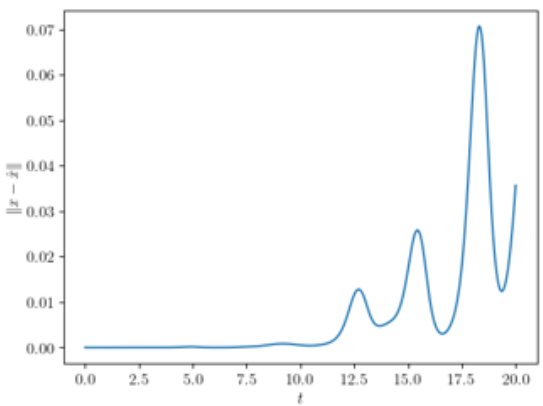
Particle flow in velocity field

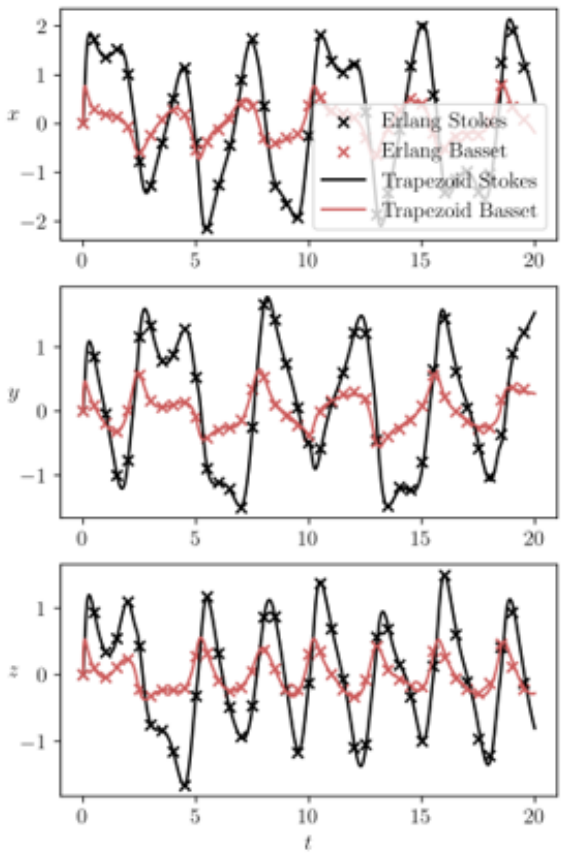
Particle subject to Stoke's drag force and Basset history force

$$\dot{x}_p = u_p,$$

$$\dot{u}_p = \underbrace{\frac{1}{\mathrm{St}}\mathcal{F}(a)(u_p - u_f)}_{\text{Nonlinear drag}} + \underbrace{C\int_0^t \frac{1}{\sqrt{t-s}}(\dot{u}_p - \dot{u}_f)\,\mathrm{d}s}_{\text{Basset history force}}$$







Collaboration with PhD candidate Zejian You, Asst. Prof. Qi Wang, and Prof. Gustaaf Jacobs from San Diego State University.

Bibliography I

[1] T. K. S. Ritschel, "Numerical optimal control for distributed delay differential equations: A simultaneous approach based on linearization of the delayed variables," in *Proceedings of the 2025 European Control Conference (ECC'25)*, (Thessaloniki, Greece), pp. 759–764, June 2025.