# Stochastic Models and Optimal Control of Epidemics under Partial Information

Ralf Wunderlich

Brandenburg University of Technology Cottbus-Senftenberg, Germany

Joint work with

Florent Ouabo Kamkumo, Ibrahim Mbouandi Njiasse (Cottbus)

Supported by

DAAD Deutscher Akademischer Austauschdienst
German Academic Exchange Service

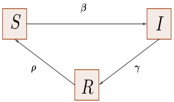
DFG Deutsche Forschungsgemeinschaft

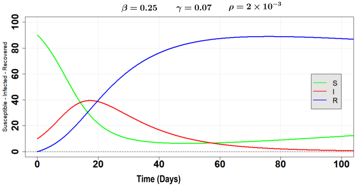
WU Vienna, Institute for Statistics and Mathematics
Research Seminar, January 19, 2024

# Introduction

## SIRS Model (Kermack and McKendrick (1927))

$$\begin{aligned}
\dot{S}(t) &= -\beta S(t)\frac{I(t)}{N} + \rho R(t) && \text{Susceptible}\\
\dot{I}(t) &= \beta S(t)\frac{I(t)}{N} - \gamma I(t) && \text{Infected}\\
\dot{R}(t) &= \gamma I(t) - \rho R(t) && \text{Recovered}
\end{aligned}$$




- System of ODEs for $X = (S, I, R)^\top : \dot{X}(t) = F(X(t))$
- Basic reproduction number $\mathcal{R}_0 = \dfrac{\beta}{\gamma}$

  Effective reproduction number $\mathcal{R}(t) = \dfrac{\beta}{\gamma}\dfrac{S(t)}{N}$

# ODE Models: Properties and Limitations

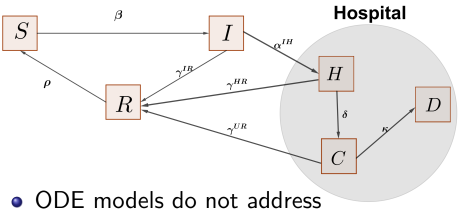
- Description and prediction of relative subpopulation sizes and
  "average" absolute subpopulation sizes

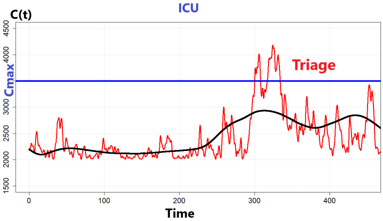
  for large total populations size $N$
- No information about deviations from the average.

  Interesting for absolute subpopulation sizes and models with

  small and moderate total population size $N$,

  small compartments such as hospitals, intensive care units (ICU).




- ODE models do not address

  uncertain parameters and initial conditions,

  forecast uncertainties,

  not directly observable subpopulation sizes

  (partial information, statistical learning of dark figures, nowcast uncertainties).

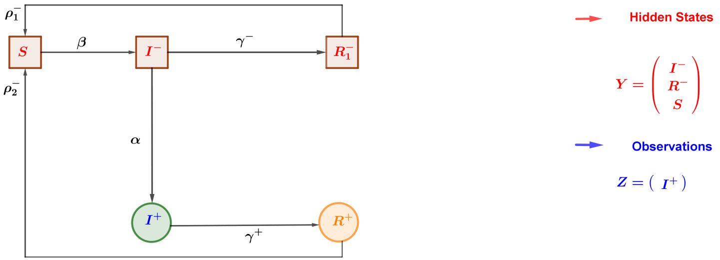
# COVID-19 Model with Partial Information

$I^-$ Infected, non-detected

$R^-$ Recovered, non-detected

S Susceptible

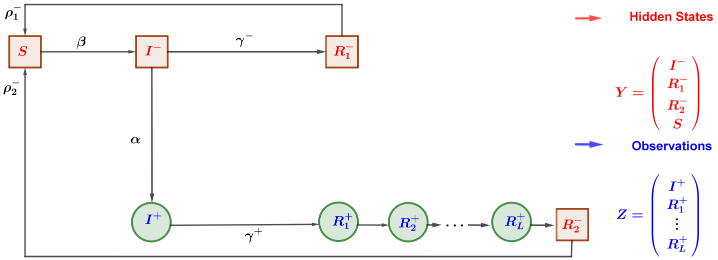
$I^+$ Infected, detected




Charpentier et al. (2020), Meyer-Hermann et al. (2021)

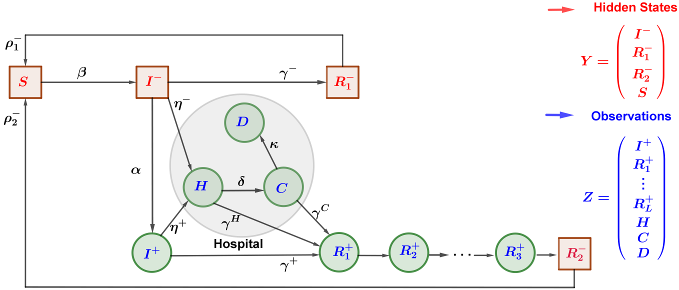
# COVID-19 Model with Partial Information

| | | | |
|---|---|---|---|
| $I^-$ | Infected, non-detected | $I^+$ | Infected, detected |
| $R_1^- / R_2^-$ | Recovered, non-detected/fading immunity | $R_1^+, \ldots, R_L^+$ | Recovered, detected |
| S | Susceptible | | |



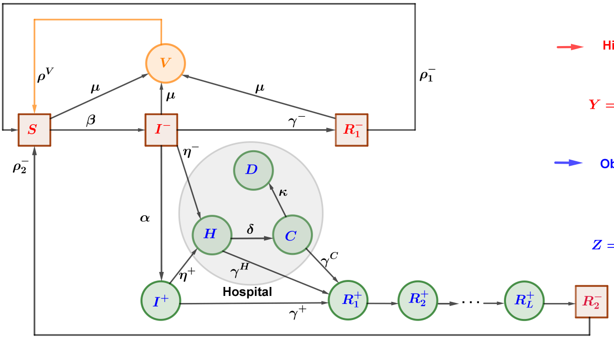
# COVID-19 Model with Hospital

| | | | |
|---|---|---|---|
| $I^-$ | Infected, non-detected | $I^+$ | Infected, detected |
| $R_1^- / R_2^-$ | Recovered, non-detected/fading immunity | $R_1^+, \ldots, R_L^+$ | Recovered, detected |
| S | Susceptible | $H$ | Hospitalized |
| | | $C$ | Intensive Care Unit |
| | | $D$ | Death |



# COVID-19 Model with Hospital and Vaccination

| | | | |
|---|---|---|---|
| $I^-$ | Infected, non-detected | $I^+$ | Infected, detected |
| $R_1^-/R_2^-$ | Recovered, non-detected/fading immunity | $R_1^+, \ldots, R_L^+$ | Recovered, detected |
| S | Susceptible | $H$ | Hospitalized |
| | | $C$ | Intensive Care Unit |
| | | $D$ | Death |

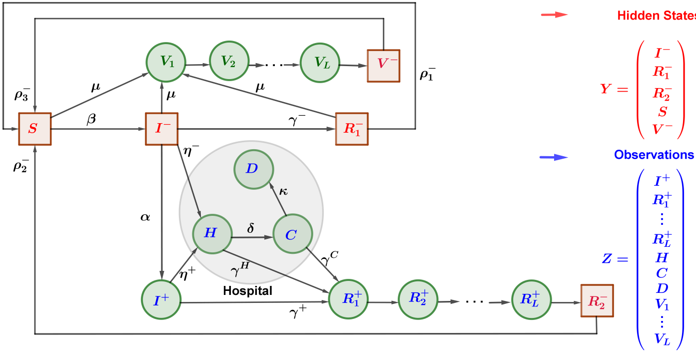



$$Y = \begin{pmatrix} I^- \\ R_1^- \\ R_2^- \\ S \\ V \end{pmatrix}$$

$$Z = \begin{pmatrix} I^+ \\ R_1^+ \\ \vdots \\ R_L^+ \\ H \\ C \\ D \end{pmatrix}$$

# COVID-19 Model with Hospital and Vaccination

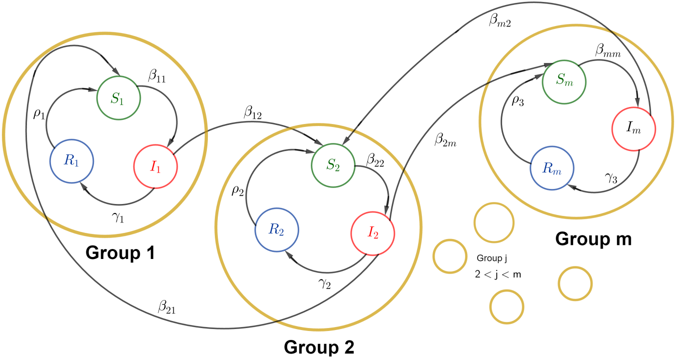
| | | | |
|---|---|---|---|
| $I^-$ | Infected, non-detected | $I^+$ | Infected, detected |
| $R_1^- / R_2^-$ | Recovered, non-detected/fading immunity | $R_1^+, \dots, R_1^+$ | Recovered detected |
| S | Susceptible | $H$ | Hospitalized |
| $V^-$ | Vaccinated, fading immunity | $C$ | Intensive Care Unint |
| | | $D$ | Death |
| | | $V_1, \dots, V_L$ | Vaccinated |




Parameters $\beta$, $\alpha$, $\mu$ may be time-dependent and controlled

# Multi-Group Models

Models with several regions, age groups, vaccination states, ...

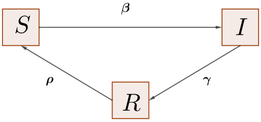


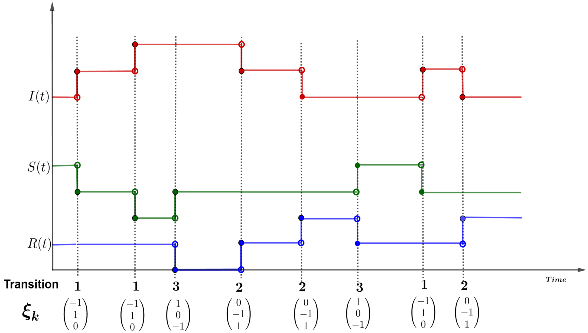
# Microscopic Stochastic Epidemic Models

- Similar to models of chemical reaction networks (Anderson, Kurtz (2011))
- Divide population of size $N$ into $d$ compartments
- $X_i(t) \in \{0, \ldots, N\}$ absolute size of subpopulation in compartment $i = 1, \ldots, d$
- $\overline{X}_i(t) = \frac{1}{N} X_i(t) \in [0, 1]$ relative size of subpopulation
- State $X = X(t) = (X_1, \ldots, X_d)^\top$ (or $\overline{X} = (\overline{X}_1, \ldots, \overline{X}_d)^\top$ )
- Individuals may undergo $K \in \mathbb{N}$ transitions between compartments
- Transition vectors $\xi_k = \Delta X(t) = X(t) - X(t-) \in \mathbb{Z}^d$ if the transition $k$ occurs at time $t$, $k = 1, \ldots, K$, typical entries of $\xi_k$ are $-1, 0, +1$
- Counting processes $M_k(t)$ : number of transition $k$ in $[0, t]$

$$X(t) = X(0) + \sum_{k=1}^{K} \xi_k M_k(t)$$

# Example: Stochastic SIRS Model




$$d = 3, \quad X = (S, I, R)^\top, \quad K = 3 \text{ transitions}$$

| k | Transition | Transition vectors $\xi_k$ | intensity $\lambda_i(x)$ |
|---|---|---|---|
| 1 | Infection of susceptible | $(-1, \quad 1, \quad 0)^\top$ | $\beta S \frac{I}{N} = \beta x_1 \frac{x_2}{N}$ |
| 2 | Recovering of infected | $(\quad 0, -1, \quad 1)^\top$ | $\gamma I = \gamma x_2$ |
| 3 | Losing immunity | $(\quad 1, \quad 0, -1)^\top$ | $\rho R = \rho x_3$ |

# Microscopic Stochastic Epidemic Models (cont.)

Recall: $X^N(t) = X^N(0) + \sum_{k=1}^{K} \xi_k M_k(t)$ state for population of size $N$

Assume $\mathbb{P}(M_k(t+\Delta t) - M_k(t) = 1 | X^N(t)) = \lambda_k(X^N(t))\Delta t + o(\Delta t)$

Describe counting processes $M_k$ by independent Poisson processes

## Continuous–Time Markov Chain (CTMC)

$$X^N(t) = X^N(0) + \sum_{k=1}^{K} \xi_k \Pi_k\Big( \int_0^t \lambda_k(X^N(s))ds \Big)$$

where $\Pi_1, \ldots, \Pi_K$ are independent standard Poisson processes

State-dependent intensities $\lambda_k = \lambda_k(X^N(s))$, for $k = 1, \cdots, K$

Assume scaling property $\lambda_k(x) = \lambda_k^N(x) = N\nu_k(N^{-1}x)$
where $\nu_k(z)$ is the intensity
in terms of relative subpopulation size $z = N^{-1}x$
and independent of $N$

Intensities may depend on time: $\lambda_k = \lambda_k(t, x)$

# Macroscopic Models & Large Population Limit

## Law of Large Numbers for Poisson Process

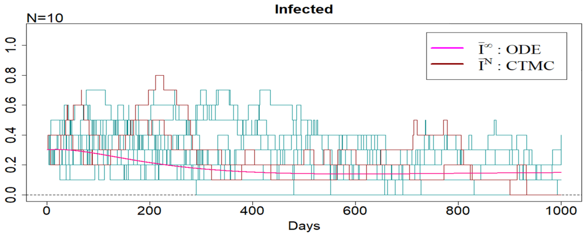
$$\left|\frac{1}{N}\Pi(Nu) - u\right| \xrightarrow[N\to\infty]{} 0 \quad \text{a.s., uniformly} \quad \text{for} \quad \text{all } u \leq u_0$$

implies $\overline{X}^N(t) \xrightarrow[N\to\infty]{} \overline{X}^\infty(t)$ uniformly for all $t \leq T$

$\overline{X}^\infty$ satisfies ODE Anderson & Kurtz (2011), Britton & Pardoux (2018)

## Large Population Limit for Relative Subpopulation Sizes

$$\frac{d}{dt}\overline{X}^\infty(t) = \overline{F}(t, \overline{X}^\infty(t)) \quad \textit{with} \quad \overline{F}(t,z) = \sum_{k=1}^{K} \xi_k \nu_k(t,z)$$



# Macroscopic Models & Large Population Limit

## Law of Large Numbers for Poisson Process

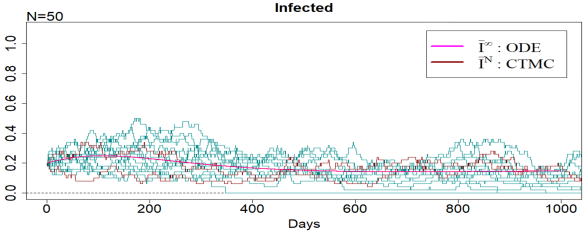
$$\left|\frac{1}{N}\Pi(Nu) - u\right| \xrightarrow[N\to\infty]{} 0 \quad \text{a.s., uniformly} \quad \text{for} \quad \text{all } u \leq u_0$$

implies $\overline{X}^N(t) \xrightarrow[N\to\infty]{} \overline{X}^\infty(t)$ uniformly for all $t \leq T$

$\overline{X}^\infty$ satisfies ODE Anderson & Kurtz (2011), Britton & Pardoux (2018)

## Large Population Limit for Relative Subpopulation Sizes

$$\frac{d}{dt}\overline{X}^\infty(t) = \overline{F}(t, \overline{X}^\infty(t)) \quad \textit{with} \quad \overline{F}(t,z) = \sum_{k=1}^{K} \xi_k \nu_k(t,z)$$



# Macroscopic Models & Large Population Limit

## Law of Large Numbers for Poisson Process

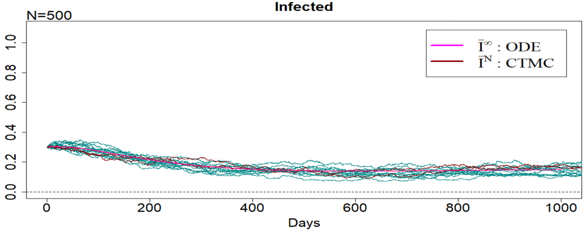
$$\left|\frac{1}{N}\Pi(Nu) - u\right| \xrightarrow[N\to\infty]{} 0 \quad \text{a.s., uniformly for all } u \leq u_0$$

implies $\overline{X}^N(t) \xrightarrow[N\to\infty]{} \overline{X}^\infty(t)$ uniformly for all $t \leq T$

$\overline{X}^\infty$ satisfies ODE Anderson & Kurtz (2011), Britton & Pardoux (2018)

## Large Population Limit for Relative Subpopulation Sizes

$$\frac{d}{dt}\overline{X}^\infty(t) = \overline{F}(t, \overline{X}^\infty(t)) \quad \textit{with} \quad \overline{F}(t,z) = \sum_{k=1}^{K} \xi_k \nu_k(t,z)$$



# Macroscopic Models & Large Population Limit

## Law of Large Numbers for Poisson Process

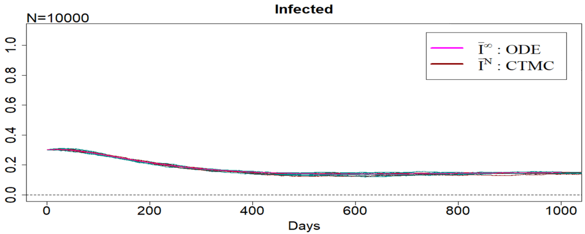
$$\left|\frac{1}{N}\Pi(Nu) - u\right| \xrightarrow[N\to\infty]{} 0 \quad \text{a.s., uniformly} \quad \text{for} \quad \text{all } u \le u_0$$

implies $\overline{X}^N(t) \xrightarrow[N\to\infty]{} \overline{X}^\infty(t)$ uniformly for all $t \le T$

$\overline{X}^\infty$ satisfies ODE Anderson & Kurtz (2011), Britton & Pardoux (2018)

## Large Population Limit for Relative Subpopulation Sizes

$$\frac{d}{dt}\overline{X}^\infty(t) = \overline{F}(t, \overline{X}^\infty(t)) \quad \textit{with} \quad \overline{F}(t,z) = \sum_{k=1}^{K} \xi_k \nu_k(t,z)$$



# Macroscopic Models & Diffusion Approximation

## Central Limit Theorem for Poisson Process

$$\frac{1}{\sqrt{N}}(\Pi(Nu) - Nu) \underset{N\to\infty}{\Longrightarrow} W(u) \quad \text{Brownian motion}$$

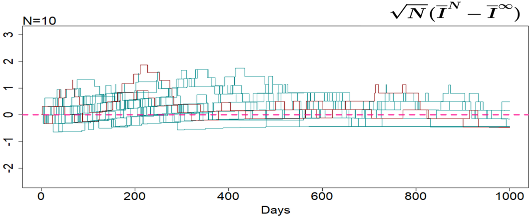
Scaling property $W(\int_0^t a(s)ds) \overset{d}{=} \int_0^t \sqrt{a(s)}dW(s)$

## Diffusion Approximation of $X^N$ and $\overline{X}^N$ by SDEs

$$\begin{array}{llll} \text{Absolute} & dX^D(t) & = F(t, X^D(t))dt & + \sigma(t, X^D(t))dW(t) \\ \text{Relative} & d\overline{X}^D(t) & = \overline{F}(t, \overline{X}^D(t))dt & + \frac{1}{\sqrt{N}}\overline{\sigma}(t, \overline{X}^D(t))dW(t) \end{array}$$

with $F(t,x) = \sum_{k=1}^{K} \xi_k \lambda_k(t,x)$, $\sigma(t,x) = (\xi_1\sqrt{\lambda_1(t,x)}, \dots, \xi_K\sqrt{\lambda_K(t,x)})$

$\overline{\sigma}(t,z) = (\xi_1\sqrt{\nu_1(t,z)}, \dots, \xi_K\sqrt{\nu_K(t,z)})$



# Macroscopic Models & Diffusion Approximation

## Central Limit Theorem for Poisson Process

$$\frac{1}{\sqrt{N}}(\Pi(Nu) - Nu) \xrightarrow[N\to\infty]{} W(u) \quad \text{Brownian motion}$$

Scaling property $W(\int_0^t a(s)ds) \stackrel{d}{=} \int_0^t \sqrt{a(s)}dW(s)$

## Diffusion Approximation of $X^N$ and $\overline{X}^N$ by SDEs

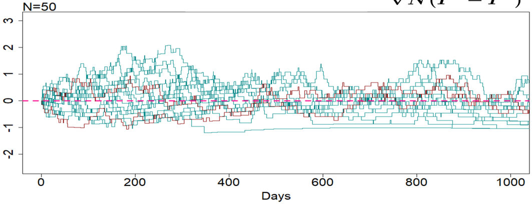
Absolute $dX^D(t) = F(t, X^D(t))dt + \sigma(t, X^D(t))dW(t)$

Relative $d\overline{X}^D(t) = \overline{F}(t, \overline{X}^D(t))dt + \frac{1}{\sqrt{N}}\overline{\sigma}(t, \overline{X}^D(t))dW(t)$

with $F(t,x) = \sum_{k=1}^{K} \xi_k \lambda_k(t,x)$, $\sigma(t,x) = (\xi_1\sqrt{\lambda_1(t,x)}, \ldots, \xi_K\sqrt{\lambda_K(t,x)})$

$\overline{\sigma}(t,z) = (\xi_1\sqrt{\nu_1(t,z)}, \ldots, \xi_K\sqrt{\nu_K(t,z)})$

$\sqrt{N}(\overline{I}^N - \overline{I}^\infty)$



# Macroscopic Models & Diffusion Approximation

## Central Limit Theorem for Poisson Process

$$\frac{1}{\sqrt{N}}(\Pi(Nu) - Nu) \xrightarrow[N\to\infty]{} W(u) \quad \text{Brownian motion}$$

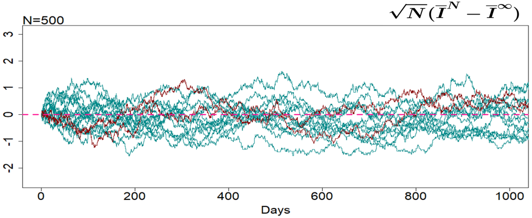
Scaling property $W\left(\int_0^t a(s)ds\right) \overset{d}{=} \int_0^t \sqrt{a(s)}dW(s)$

## Diffusion Approximation of $X^N$ and $\overline{X}^N$ by SDEs

$$\begin{aligned} \text{Absolute} \quad dX^D(t) &= F(t, X^D(t))dt + \sigma(t, X^D(t))dW(t) \\ \text{Relative} \quad d\overline{X}^D(t) &= \overline{F}(t, \overline{X}^D(t))dt + \frac{1}{\sqrt{N}}\overline{\sigma}(t, \overline{X}^D(t))dW(t) \end{aligned}$$

with $F(t,x) = \sum_{k=1}^{K} \xi_k \lambda_k(t,x)$, $\sigma(t,x) = (\xi_1\sqrt{\lambda_1(t,x)}, \dots, \xi_K\sqrt{\lambda_K(t,x)})$

$\overline{\sigma}(t,z) = (\xi_1\sqrt{\nu_1(t,z)}, \dots, \xi_K\sqrt{\nu_K(t,z)})$



# Macroscopic Models & Diffusion Approximation

## Central Limit Theorem for Poisson Process

$$\frac{1}{\sqrt{N}}(\Pi(Nu) - Nu) \xrightarrow[N\to\infty]{} W(u) \quad \text{Brownian motion}$$

Scaling property $W\left(\int_0^t a(s)ds\right) \stackrel{d}{=} \int_0^t \sqrt{a(s)}dW(s)$

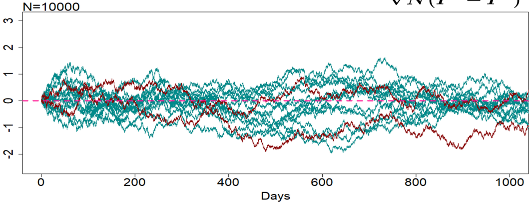
## Diffusion Approximation of $X^N$ and $\overline{X}^N$ by SDEs

$$\begin{aligned} \text{Absolute} \quad dX^D(t) &= F(t, X^D(t))dt + \sigma(t, X^D(t))dW(t) \\ \text{Relative} \quad d\overline{X}^D(t) &= \overline{F}(t, \overline{X}^D(t))dt + \frac{1}{\sqrt{N}}\overline{\sigma}(t, \overline{X}^D(t))dW(t) \end{aligned}$$

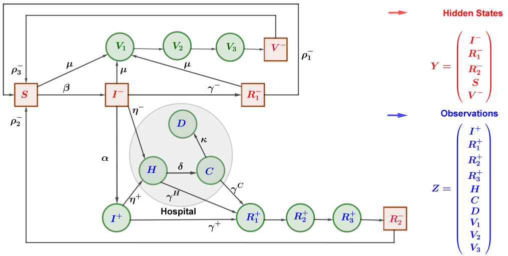
with $F(t,x) = \sum_{k=1}^{K} \xi_k \lambda_k(t,x)$, $\sigma(t,x) = (\xi_1\sqrt{\lambda_1(t,x)}, \dots, \xi_K\sqrt{\lambda_K(t,x)})$

$\overline{\sigma}(t,z) = (\xi_1\sqrt{\nu_1(t,z)}, \dots, \xi_K\sqrt{\nu_K(t,z)})$

$\sqrt{N}(\overline{I}^N - \overline{I}^\infty)$



# Diffusion Approximation of COVID-19 Model




$$X = \begin{pmatrix} Y \\ Z \end{pmatrix}$$

$$\begin{aligned} dY &= \quad \bar{f}(t,Y,Z)dt + \bar{\sigma}(t,Y,Z)dW^1 + \bar{g}(t,Y,Z)dW^2 \quad \text{hidden state} \\ dZ &= \underbrace{[\bar{h}_0(t,Z) + \bar{h}_1(t,Z)Y]}_{\text{linear in } Y} dt + \bar{\ell}(t,Y,Z)dW^2 \quad \text{observation} \end{aligned}$$

Coefficients $\bar{f}$, $\bar{\sigma}$, $\bar{g}$, $\bar{\ell}$ are non-linear in the hidden state $Y$

Time discretization, $t_n = n\Delta t$, $n = 0, 1, \ldots$

$$\begin{aligned} Y_{n+1} &= Y_n + \quad f(n, Y_n, Z_n) + \sigma(n, Y_n, Z_n)\mathcal{E}^1_{n+1} + g(n, Y_n, Z_n)\mathcal{E}^2_{n+1} \\ Z_{n+1} &= Z_n + [h_0(n, Z_n) + h_1(n, Z_n)Y_n] + \ell(n, Y_n, Z_n)\mathcal{E}^2_{n+1} \end{aligned}$$

$(\mathcal{E}^1_n), (\mathcal{E}^2_n)$ independent sequences of i.i.d. $\mathcal{N}(0, \mathbb{1})$ random vectors

Given the observations of $Z_n$ we want to estimate hidden state $Y_n$

# Filtering Problem

Decompose state vector $X = (Y, Z)^\top$ into

Y: hidden (non-observable) state

Z: observation

Given observations $Z_k$ for $k = 0, \dots, n$ and

$\mathcal{F}_0^I$ initial information about distribution of $Y_0$

Mean-square optimal estimate of $Y_n$ given $\mathcal{F}_n^Z = \sigma\{Z_k, k = 0, \dots, n\} \vee \mathcal{F}_0^I$ is

### Conditional Mean

$$M_n = \mathbb{E}\left[Y_n | \mathcal{F}_n^Z\right]$$

Measure of estimation error

### Conditional Covariance

$$Q_n := \mathrm{Var}(Y_n | \mathcal{F}_n^Z) = \mathbb{E}[(Y_n - M_n)(Y_n - M_n)^\top | \mathcal{F}_n^Z]$$

Initial estimates

$$M_0 = m_0 = \mathbb{E}[Y_0 | \mathcal{F}_0^Z] \quad \text{and}$$

$$Q_0 = q_0 = \mathrm{Var}(Y_0 | \mathcal{F}_0^Z)$$

# Kalman Filter for Conditionally Gaussian Sequences

$$Y_{n+1} = [f_0 + f_1 Y_n] + \sigma \mathcal{E}^1_{n+1} + g\, \mathcal{E}^2_{n+1} \qquad \text{hidden state/signal}$$

$$Z_{n+1} = [h_0 + h_1 Y_n] + \ell\, \mathcal{E}^2_{n+1} \qquad \text{observation}$$

$(\mathcal{E}^1_n), (\mathcal{E}^2_n)$ independent sequences of i.i.d. $\mathcal{N}(0, \mathbb{1})$ random vectors

Coefficients $a = f_0, f_1, h_0, h_1, \sigma, g, \ell$ are of the form $a = a(n, \mathcal{Z}_n)$

may depend on time $n$

and also on whole observation path $\mathcal{Z}_n = (Z_k)_{k \le n}$ up to time $n$

**Theorem (Liptser & Shiryaev (2001), Theorem 13.4)**

*Under technical assumptions the conditional distribution of $Y_n$ given $\mathcal{F}^Z_n$ is $\mathcal{N}(M_n, Q_n)$ (Gaussian).*

*$M_n$ and $Q_n$ are defined by the following recursions driven by the observations*

$$M_{n+1} = [f_0 + f_1 M_n] + \left[g\ell^\top + f_1 Q_n h_1^\top\right] \left[\ell\ell^\top + h_1 Q_n h_1^\top\right]^+ \left[Z_{n+1} - \left(h_0 + h_1 M_n\right)\right]$$

$$Q_{n+1} = -\left[g\ell^\top + f_1 Q_n h_1^\top\right] \left[\ell\ell^\top + h_1 Q_n h_1^\top\right]^+ \left[g\ell^\top + f_1 Q_n h_1^\top\right]^\top + f_1 Q_n f_1^\top + \sigma\sigma^T$$

*with initial values $M_0 = m_0$, $Q_0 = q_0$ and $Y_0 \sim \mathcal{N}(m_0, q_0)$.*

*($[A]^+$ denotes the pseudoinverse of $A$)*

Note that all coefficients may depend on time $n$ and the observation path $\mathcal{Z}_n$.

# Extended Kalman Filter

$$\begin{aligned} Y_{n+1} &= Y_n + f(n, Y_n, Z_n) + \sigma(n, Y_n, Z_n)\mathcal{E}^1_{n+1} + g(n, Y_n, Z_n)\mathcal{E}^2_{n+1} \\ Z_{n+1} &= Z_n + [h_0(n, Z_n) + h_1(n, Z_n)Y_n] + \ell(n, Y_n, Z_n)\mathcal{E}^2_{n+1} \end{aligned}$$

Drift coefficient $f$ non-linear w.r.t. signal $Y$

Diffusion coefficients $\sigma, g, \ell$ may also depend on signal $Y$

Idea: Gelb (1974), Pardoux (1991), Bain, Crisan (2009)

1. Linearize drift coefficient $f$ by Taylor expansion around a "suitable" $\overline{Y}_n$
2. Substitute signal $Y$ by $\overline{Y}_n$ in diffusion coefficients $\sigma, g, \ell$

### Approximation by Conditional Gaussian Sequences

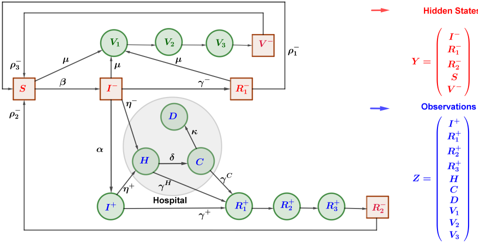
$$\begin{aligned} \widetilde{Y}_{n+1} &= \widetilde{Y}_n + f(n, \overline{Y}_n, \widetilde{Z}_n) + \frac{\partial f}{\partial y}(n, \overline{Y}_n, \widetilde{Z}_n)(\widetilde{Y}_n - \overline{Y}_n) \\ &\quad + \sigma(n, \overline{Y}_n, \widetilde{Z}_n)\mathcal{E}^1_{n+1} + g(n, \overline{Y}_n, \widetilde{Z}_n)\mathcal{E}^2_{n+1} \\ \widetilde{Z}_{n+1} &= \widetilde{Z}_n + [h_0(n, \widetilde{Z}_n) + h_1(n, \widetilde{Z}_n)\widetilde{Y}_n] + \ell(n, \overline{Y}_n, \widetilde{Z}_n)\mathcal{E}^2_{n+1} \end{aligned}$$

3. Apply Kalman filter for conditional Gaussian sequences $\rightsquigarrow (M_n, Q_n)$

Extended Kalman filter linearizes around current filter estimate: $\overline{Y}_n = M_n$.

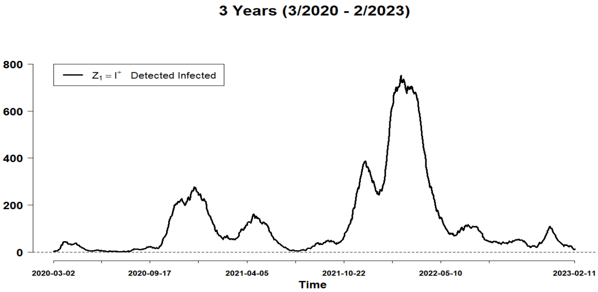
For theoretical justification and error estimates see Picard (1991)

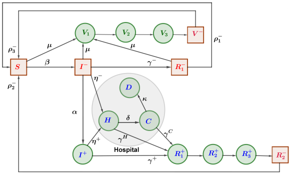
- COVID-19 model with partial information



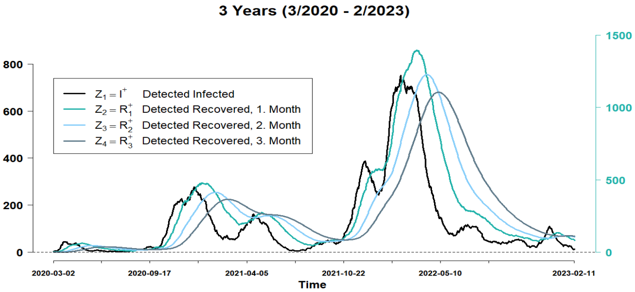

- Calibrate model parameters to real-world data set for Germany
  Time-depending $\beta$ and $\alpha$ match daily basic reproduction numbers,
  positive tests
- $T = 3$ years from March 2020 to February 2023, $\Delta t = 1$ day
- Population size $N = 100.000$
- Simulate hidden states and observations
- Compute filter estimates of hidden states based on observations
- Compare estimated and true values
  $\longrightarrow$ precision of the filter estimates

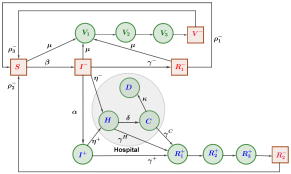
# Observations: Detected Infected



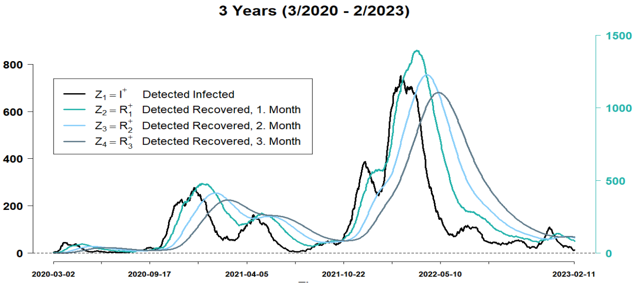



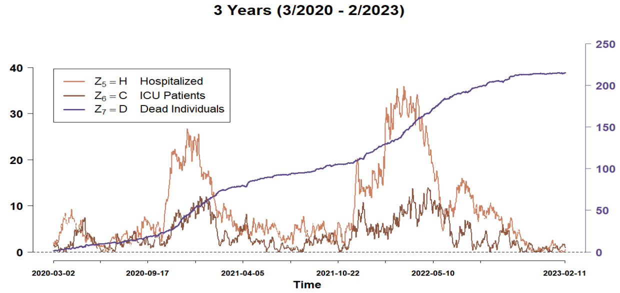
# Observations: Detected Infected & Recovered

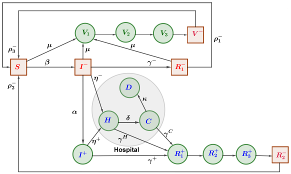


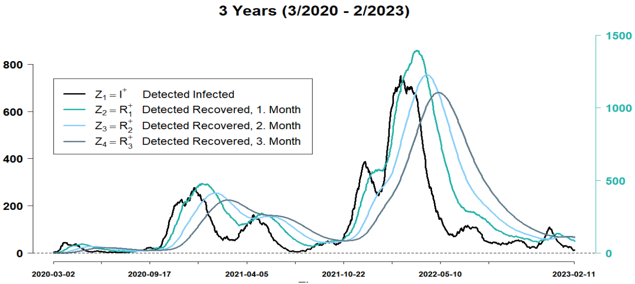
# Observations: Hospital

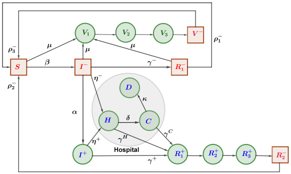


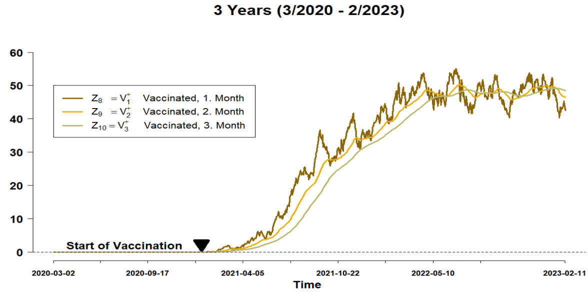



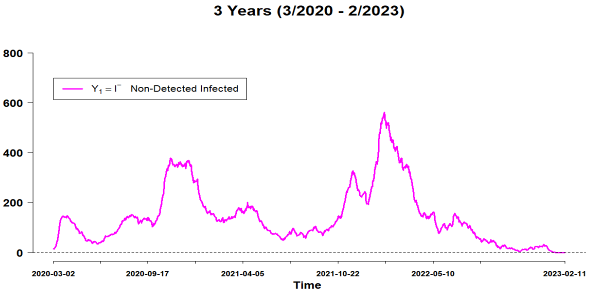
# Observations: Vaccinated

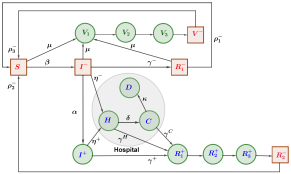


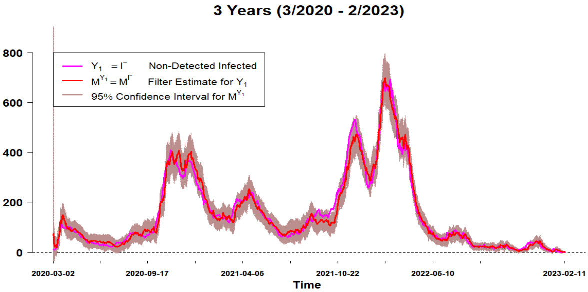



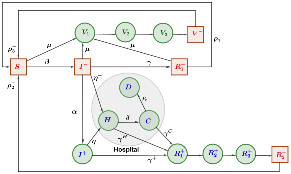
# Hidden Signal Simulation: Non-Detected Infected $I^-$



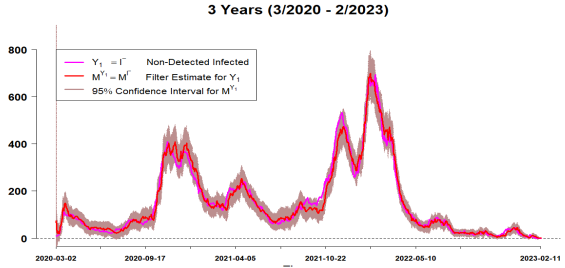



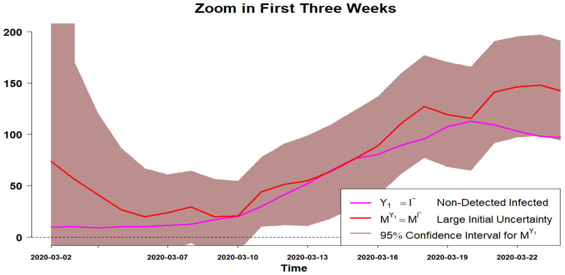
# Hidden Signal Estimation: Non-Detected Infected $I^-$

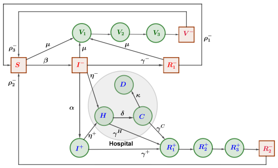


# Hidden Signal Estimation: Non-Detected Infected $I^-$

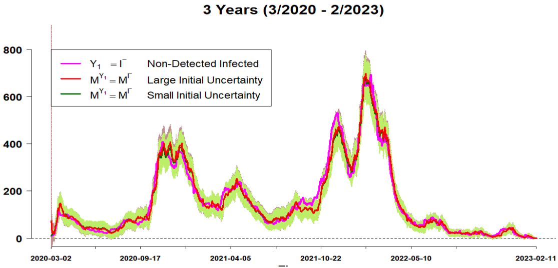


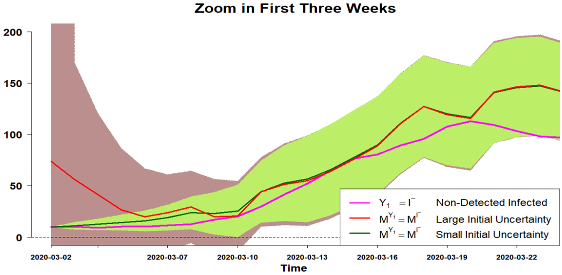



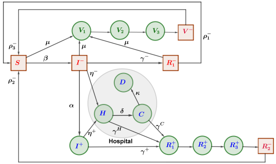




Large initial uncertainty is reduced by learning from observations

# Impact of Initial Estimate & Effect of Learning

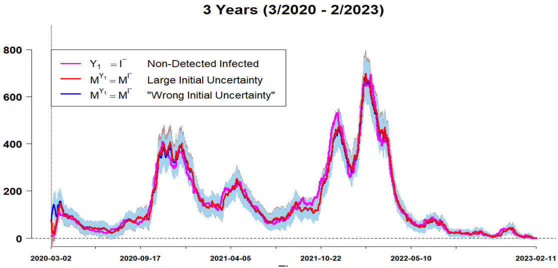


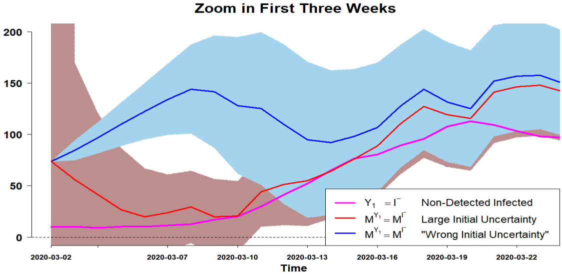



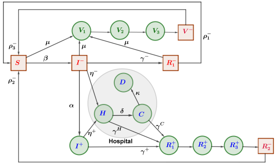




| | |
|---|---|
| Red | Large initial uncertainty is reduced by learning from observations |
| Green | Small initial uncertainty is fading out by observation noise |

# Impact of Initial Estimate & Effect of Learning




Red Large initial uncertainty is reduced by learning from observations

Blue "Wrong" initial uncertainty needs long time to be corrected

# Optimal Control Problem

For containment of an epidemics decision makers (government) try to influence the course of the epidemics by

1. Social distancing / lock-down with (relative) force $u_1 \in [0, 1]$, reduces transmission rate from $\beta$ to $(1 - u_1)\beta$
2. Tests/Diagnosis with intensity $u_2 \geq 0$
3. Vaccination with intensity $u_3 \geq 0$,

These measures have financial or social costs.

Available capacities for testing and vaccination are limited.

Aim: cost-optimal containment of the epidemics
through an appropriate mix of measures

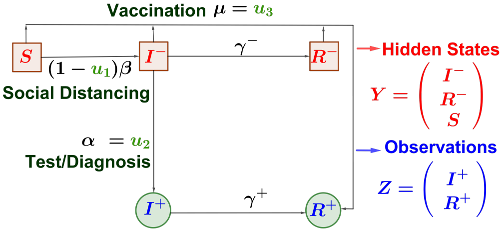
Decision-making problem under uncertainty about

future course of epidemics (forecasts),
current state of epidemics (nowcasts, dark numbers)

# Simplified Model

Model a disease with lifelong immunity after infection or vaccination

Example: measles



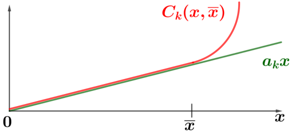
# Running and Terminal Cost

Running cost proportional to number $x$ of "affected" individuals,

Penalties if capacity threshold $\overline{x}$ is exceeded

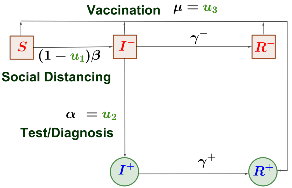
$C_k : \mathbb{R}^2 \to \mathbb{R}_+, \quad (x, \overline{x}) \mapsto C_k(x, \overline{x})$

increasing and convex functions w.r.t. $x$,




Example: $C_k(x, \overline{x}) = \begin{cases} a_k x, & x \le \overline{x} \\ a_k x + b_k(x - \overline{x})^2, & x > \overline{x} \end{cases}$

$\overline{x} = 0$: quadratic, $\overline{x} \to \infty$: linear




| | | |
|---|---|---|
| Social distancing (Lockdown) | $C_1(u_1 X^{\text{Work}}, 0)$, | $X^{\text{Work}} = N - I^+$ Charpentier et al. (2020) |
| Tests | $C_2(u_2 X^{\text{Test}}, \overline{x}^{\text{Test}})$, | $X^{\text{Test}} = N - I^+ - R^+$ |
| Vaccination | $C_3(u_3 X^{\text{Test}}, \overline{x}^{\text{Vacc}})$ | |
| Penalties | $C_4^{\pm}(I_{N_t}^{\pm}, 0)$ | for infected |

Running cost $\quad \Psi(X_n, u_n) = C_1 + C_2 + C_3 + C_4^+ + C_4^-$

Terminal cost $\quad \Phi(X_{N_t}) = C_T^+(I_{N_t}^+, 0) + C_T^-(I_{N_t}^-, 0)$

Penalties for infected at terminal time $T = N_t \Delta t$

# Performance Criterion

**Expected Aggregated Cost: Full Information, $X = (Y, Z)$**

$$\mathcal{J}^F(x;u) = \mathbb{E}\Big[\sum_{n=0}^{N_t-1} \Psi(X_n, u_n) + \Phi(X_{N_t})\Big| X_0 = x\Big]$$

Problem: $X = (Y, Z)$ depends on hidden state $Y$

Initial state $X_0 = x = (y, z)$ is not known

Decisions (control $u$) have to based on observable quantities $Z$ only

- Take conditional expectation w.r.t. $\mathcal{F}_0^Z$ (initial information)

$$\begin{aligned}\mathbb{E}[\mathcal{J}^F(X_0;u)|\mathcal{F}_0^Z] &= \mathbb{E}\Big[\sum_{n=0}^{N_t-1} \Psi(\underbrace{(Y_n, Z_n)}_{=X_n}, u_n) + \Phi(Y_{N_t}, Z_{N_t})\Big|\mathcal{F}_0^Z\Big] \\ &\overset{\substack{\text{tower law} \\ \text{Fubini}}}{=} \mathbb{E}\Big[\sum_{n=0}^{N_t-1} \mathbb{E}\Big[\Psi((Y_n, Z_n), u_n)\Big|\mathcal{F}_n^Z\Big] + \mathbb{E}\Big[\Phi(Y_{N_t}, Z_{N_t})\Big|\mathcal{F}_{N_t}^Z\Big]\Big|\mathcal{F}_0^Z\Big]\end{aligned}$$

- For $\Psi, \Phi$ linear and quadratic in $y$ conditional expectation $\mathbb{E}[\dots|\mathcal{F}_n^Z]$ can be expressed in terms of Extended Kalman filter $(M, Q)$ for hidden state $Y$.
  Recall: conditional distribution of $Y_n$ is Gaussian $\mathcal{N}(M_n, Q_n)$

**Performance Criterion: Partial Information, $X^P = (M, Q, Z)$**

$$\mathcal{J}(x^P;u) = \mathbb{E}\Big[\sum_{n=0}^{N_t-1} \Psi^P(X_n^P, u_n) + \Phi^P(X_{N_t}^P)\Big| X_0^P = x^P\Big]$$

# Optimal Control Problem with Partial Information

Replace hidden state $Y$ by filter $(M, Q)$

Rewrite dynamics of $M$ and $Z$ in terms of innovations process $(\overline{\mathcal{E}}_n)$

with $\quad \overline{\mathcal{E}}_n = ([\ell\ell^\top + h_1 Q_{n-1} h_1^\top]^+)^{1/2}(Z_n - (h_0 + h_1 M_{n-1})),$

$\quad (\overline{\mathcal{E}}_n)$ i.i.d. $\mathcal{N}(0, \mathbb{1})$ r.v.'s with $\mathcal{F}_n^Z = \mathcal{F}_n^{\overline{\mathcal{E}}} \vee \mathcal{F}_0^I$ (Liptser, Shiryaev (2001))

Treat control problem as Markov decision process (MDP) with

- state process $\quad X^P = (M, Q, Z)^\top$ taking values in state space $\mathcal{X}$
- dynamics $\quad X_{n+1}^P = \mathcal{T}(n, X_n^P, u_n, \overline{\mathcal{E}}_{n+1})$ with transition operator $\mathcal{T}$
- and Gaussian transition kernel.

$X^P = (M, Q, Z)$ is adapted to the observable filtration

$$\mathbb{F}^Z = (\mathcal{F}_n^Z)_{n \geq 0} \quad \text{with} \quad \mathcal{F}_n^Z = \sigma\{Z_k, k \leq n\} \vee \mathcal{F}_0^I$$

Admissible controls

$$\mathcal{A} = \{(u_n)_{n=0,\dots,N_t-1} \mid \mathbb{F}^Z\text{-adapted}, \ \text{integrability cond., Markov control } u_n = \widetilde{u}(n, X_n^P), \ \text{control constraints } u_n \in \mathcal{U} = [0,1] \times \mathbb{R}_+^2\}$$

Performance criterion for $n = 0, \dots, N_t$, $x = (m, q, z)^\top$ and $u \in \mathcal{A}$

$$J(n, x; u) = \mathbb{E}\Big[\sum_{k=n}^{N_t-1} \Psi^P(X_k^P, u_k) + \Phi^P(X_{N_t}^P) \Big| X_n^P = x\Big]$$

**Optimization problem**

Find $u^* \in \mathcal{A}$ such that $\quad J(n, x; u^*) = V(n, x) := \inf_{u \in \mathcal{A}} J(n, x; u)$

# Solution Using Dynamic Programming

**Bellman Equation / Dynamic Programming Equation**

$$V(n,x) = \inf_{\nu \in \mathcal{U}} \Big\{ \Psi^P(x,\nu) + \mathbb{E}_{n,x}\big[V(n+1, \underbrace{\mathcal{T}(n,x,\nu,\overline{\mathcal{E}}_{n+1})}_{=X^P_{n+1}})\big] \Big\}, \; n = 0,1,\dots,N-1$$

$$V(N_t, x) = \Phi^P(x) \qquad \text{(terminal condition)}$$

...can be solved by backward recursion

**Challenge**
- Compute $\mathbb{E}_{n,x}\big[V(n+1, \mathcal{T}(n,x,\nu,\overline{\mathcal{E}}_{n+1}))\big]$ at each time $n$ for all $x \in \mathcal{X}$!
- No closed-form expressions of the expectation are available.
- For high-dimensional state this becomes computationally intractable.
- $\rightarrow$ Curse of dimensionality
- Simplified model: dimension of state $X^P = (M, Q, Z)^\top$ is 7.

**Remedies**
- Apply quantization techniques as in Pagès (2015), Callegaro et al. (2017)
- Model order reduction: PCA for covariance matrix $Q$
- Q-Learning
- ⋮

# References


Anderson, D. F. and Kurtz, T. G. (2011),
Continuous-Time Markov Chain Models for Chemical Reaction Networks,
*In Design and analysis of biomolecular circuits*, *Springer, 3-42*.

Bain, A. and Crisan, D. (2009), Fundamentals of Stochastic Filtering, *Springer* .

Britton T., Pardoux, E. (2019),
Stochastic Epidemic Models With Inference, *Springer*.

Charpentier A., Elie R., Laurière M., and Tran V. Chí (2020),
COVID-19 Pandemic Control: Balancing Detection Policy and Lockdown Intervention Under ICU Sustainability. *Math. Model. Nat. Phenom.* (15) 57.

Gelb, A. (1974), Applied Optimal Estimation, *MIT press*.

Liptser, R. S. and Shiryaev, A. N. (2001),
*Statistics of Random Processes II: Applications,* Vol.2, *Springer*.

Pardoux, É (1991), Filtrage Non Lineaire Et Equations Aux Derivees Partielles Stochastiques Associees, In *Lecture Notes in Mathematics*, Vol. 1464, 67–163, Springer.

Picard, J. (1991), Efficiency of the Extended Kalman Filter for Nonlinear Systems with Small Noise. *SIAM Journal on Applied Mathematics*, 51, 843-885.