Epidemic renewal models and sequential Monte Carlo

Nicholas Steyn - 30 May 2025
Bioinference

UNIVERSITY OF OXFORD
DEPARTMENT OF STATISTICS
nicholas.steyn@univ.ox.ac.uk

Who is this talk for?

An intro/refresher on **SMC methods**

An intro/refresher on the **renewal model**

A (small) handful of **real-world results**

(Statistical) modeller

Public health epi

Statistical epidemiologist

Renewal Models

A simple ***model*** of an epidemic

$$C_t \sim \text{Poisson}\left(R_t \sum_{u=1}^{\omega_{max}} C_{t-u}\omega_u \right)$$

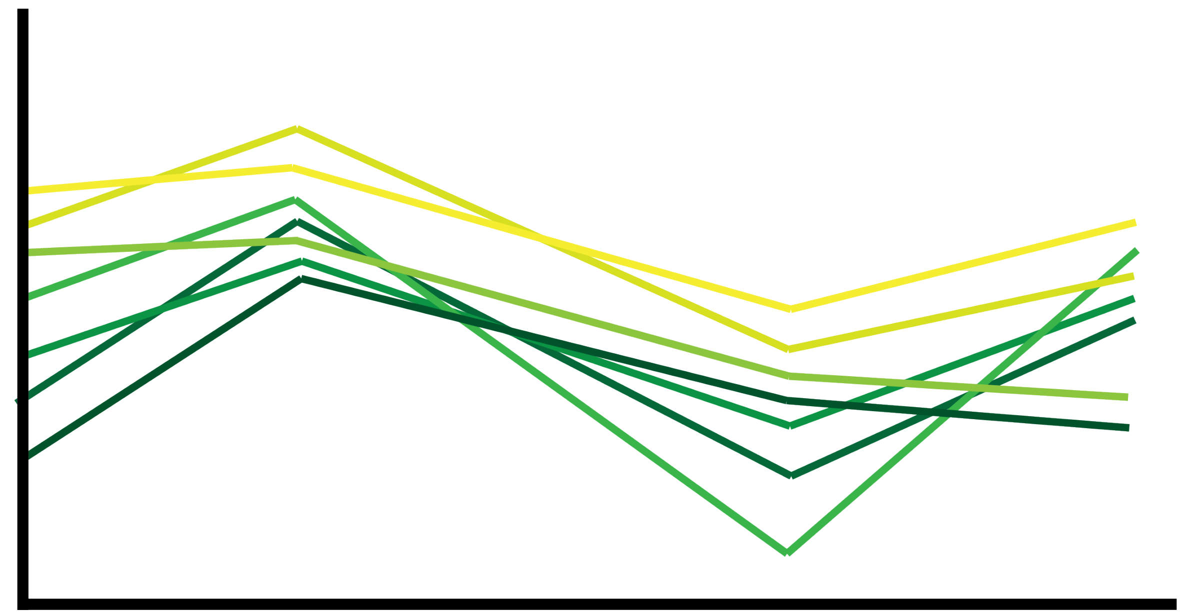
Used for reproduction number estimation

Also forecasting, elimination, effect of NPIs, etc

EpiEstim/EpiNow2/EpiFilter all use it

Sequential Monte Carlo

A ***method*** for fitting hidden-state models



Can handle any *sequential* model

Also known as “particle filters”

Very flexible!

Renewal Models

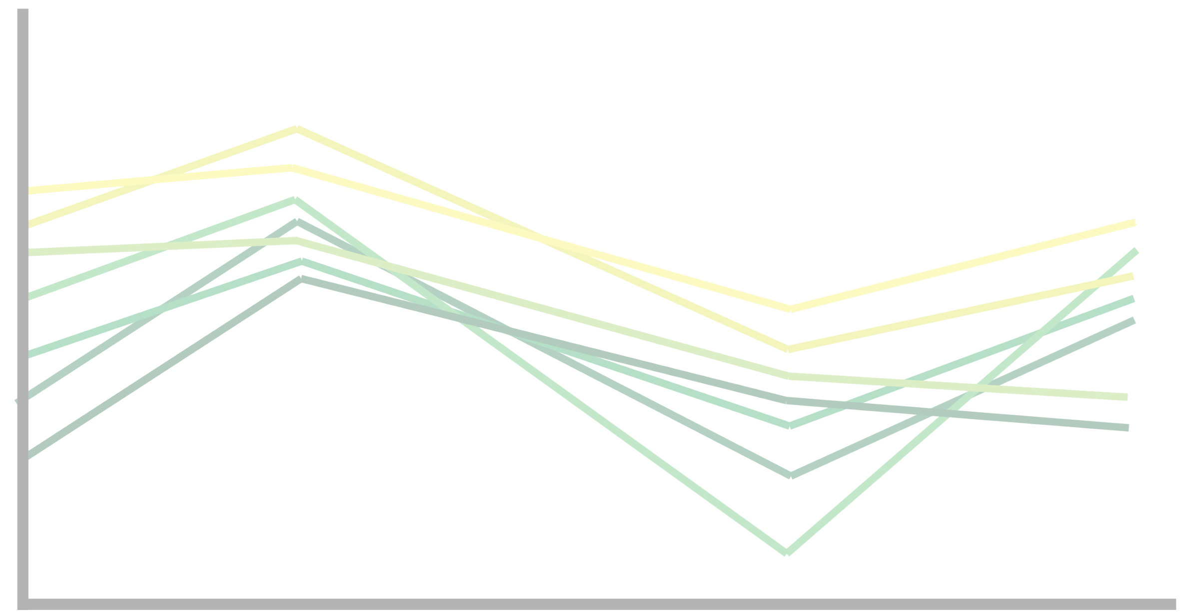
A simple ***model*** of an epidemic

$$C_t \sim \text{Poisson}\left(R_t \sum_{u=1}^{\omega_{max}} C_{t-u}\omega_u\right)$$

Used for reproduction number estimation

Also forecasting, elimination, effect of NPIs, etc

EpiEstim/EpiNow2/EpiFilter all use it



The renewal model

- Originally attributed to Euler in 1767, modelling population dynamics
- Lotka gave us the current (continuous time form) in 1907:

$$B(t) = \int_0^t B(t-a)p(a)v(a)\ da$$

births P(survival) fertility

The renewal model

- Originally attributed to Euler in 1767, modelling population dynamics

- Lotka gave us the current (continuous time form) in 1907:

$$B(t) = \int_0^t B(t-a)p(a)v(a)\ da$$

~~births~~ ~~P(survival)~~ ~~fertility~~

infections P(infectious) infectiousness

- Adapted to infections by Kermack and McKendrick (1927)*

The renewal model

- A more familiar form was introduced in (Diekmann, 1977)*:

$$I(t) = \int_0^t I(t-\tau)A(\tau)\ d\tau$$

The renewal model

- A more familiar form was introduced in (Diekmann, 1977)
- Adapted to allow for time-varying average infectiousness:

$$I(t) = \int_0^t I(t-\tau)A(t,\tau)\ d\tau$$

The renewal model

- A more familiar form was introduced in (Diekmann, 1977)
- Adapted to allow for time-varying average infectiousness, then
- (Fraser, 2007) separated infectiousness into the reproduction number and the generation time distribution*:

$$I(t) = R(t) \int_0^t I(t-\tau)\omega(\tau)\, d\tau$$

The renewal model

- A more familiar form was introduced in (Diekmann, 1977)
- Adapted to allow for time-varying average infectiousness, then
- (Fraser, 2007) separated infectiousness into the reproduction number and the generation time distribution
- and introduced the discrete-time analogue:

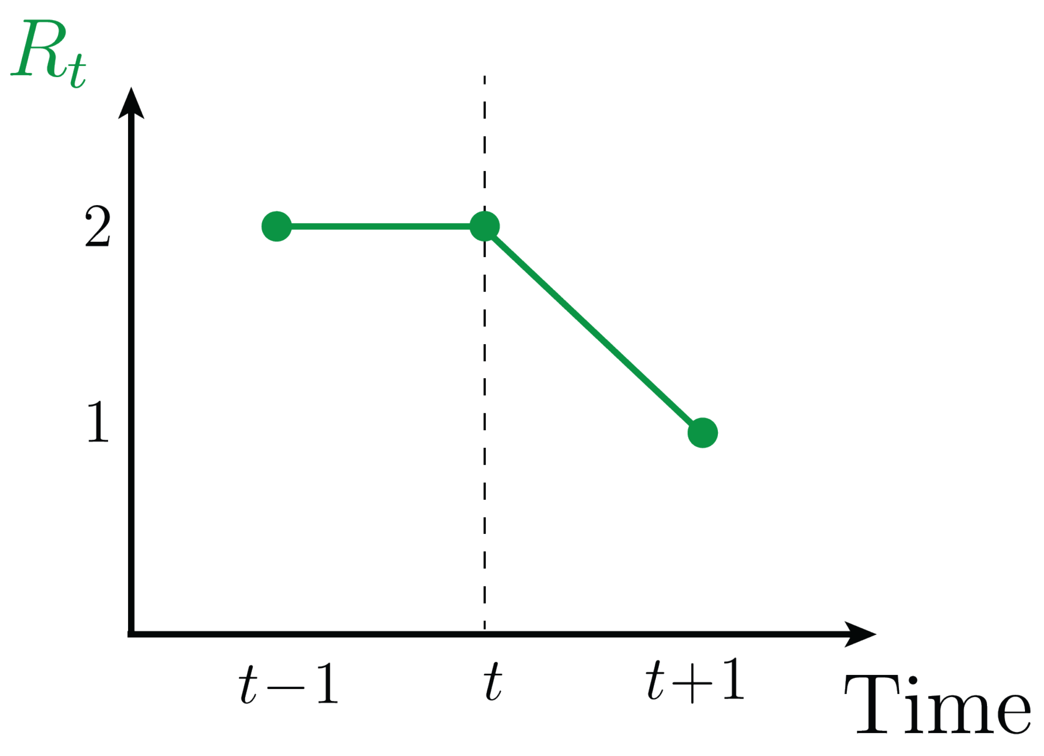
$$I_t = R_t \sum_{\tau=0}^{t} I_{t-\tau} \omega_\tau$$

- Which was then popularised by the EpiEstim software package (Cori et al, 2013)

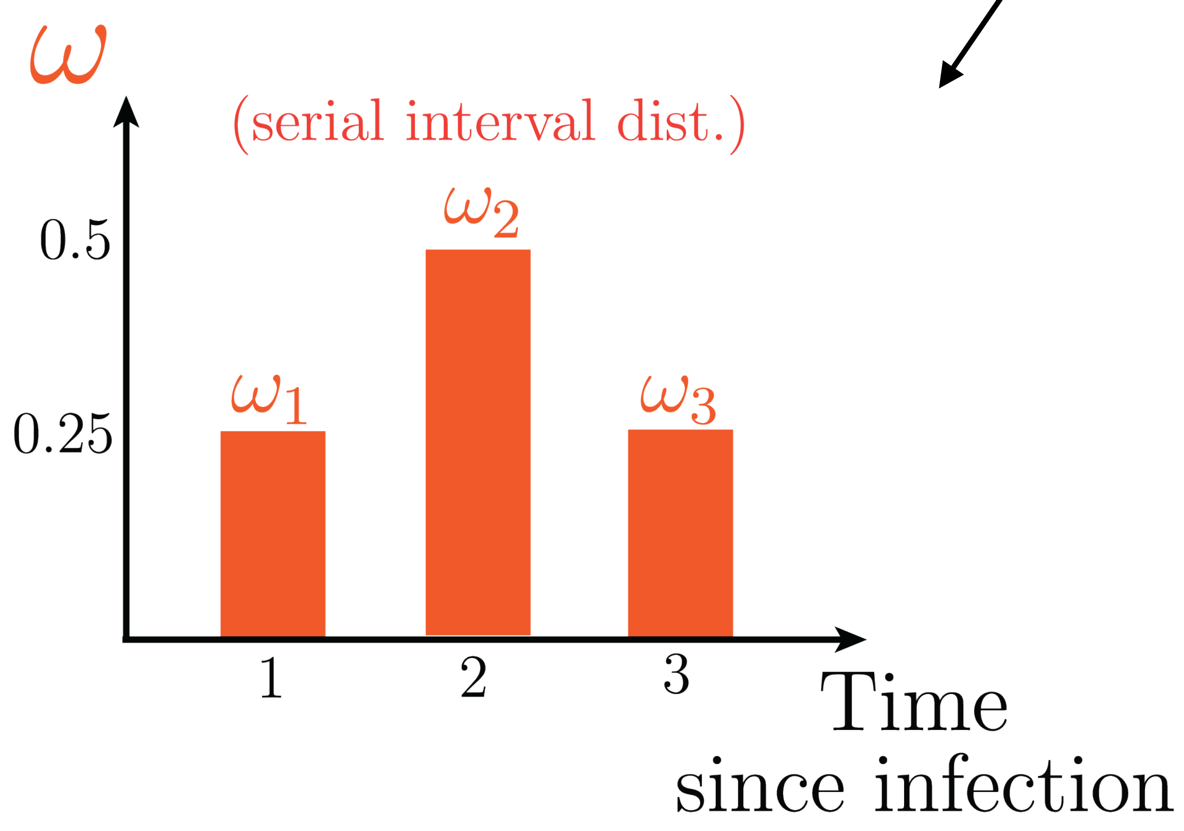
The renewal model

A toy example

Reproduction number over time

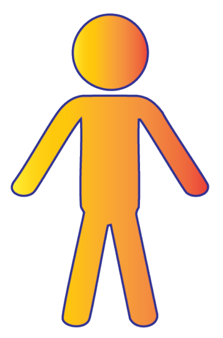


A **serial interval distribution**

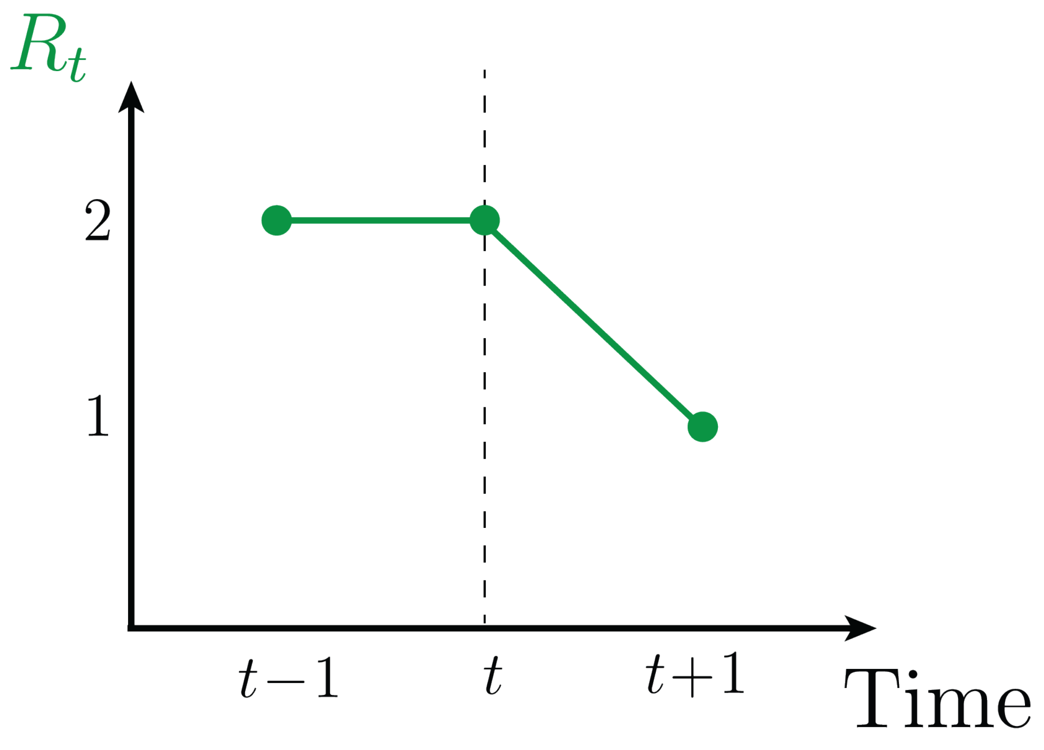


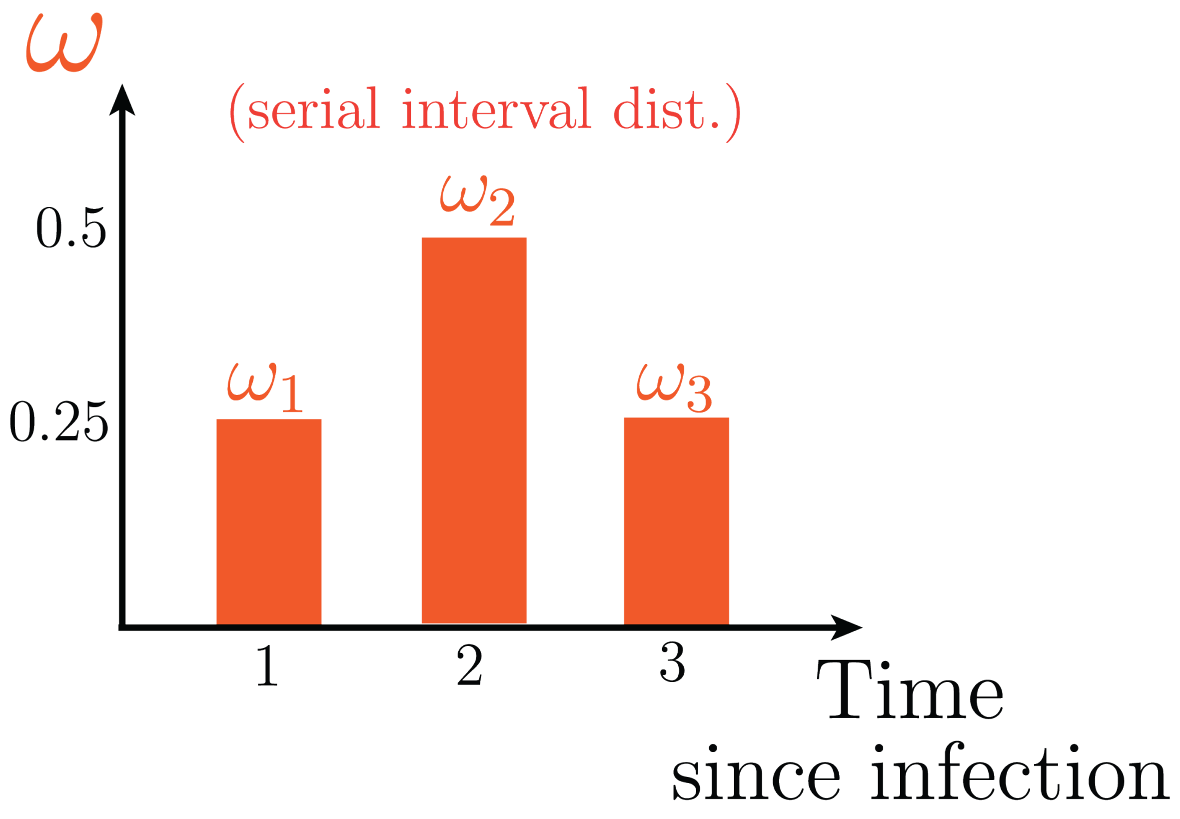
The renewal model

Someone infected two days ago…

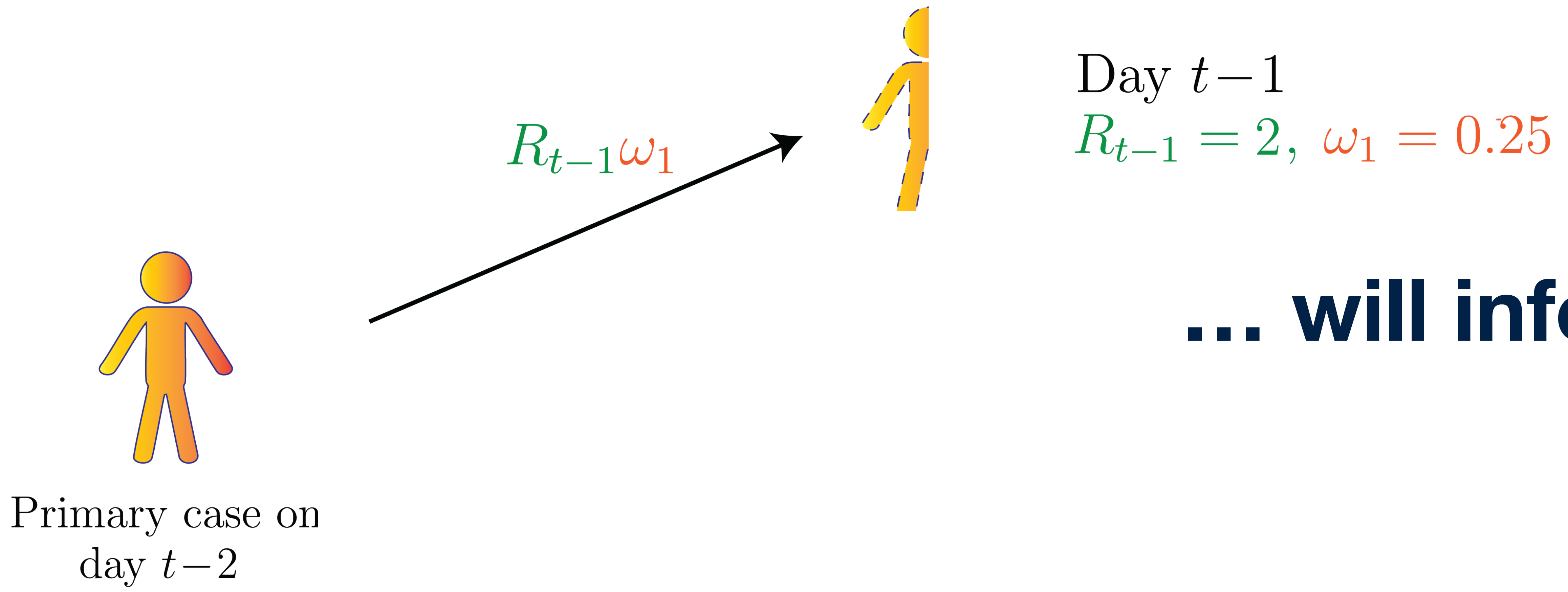


Primary case on day $t-2$

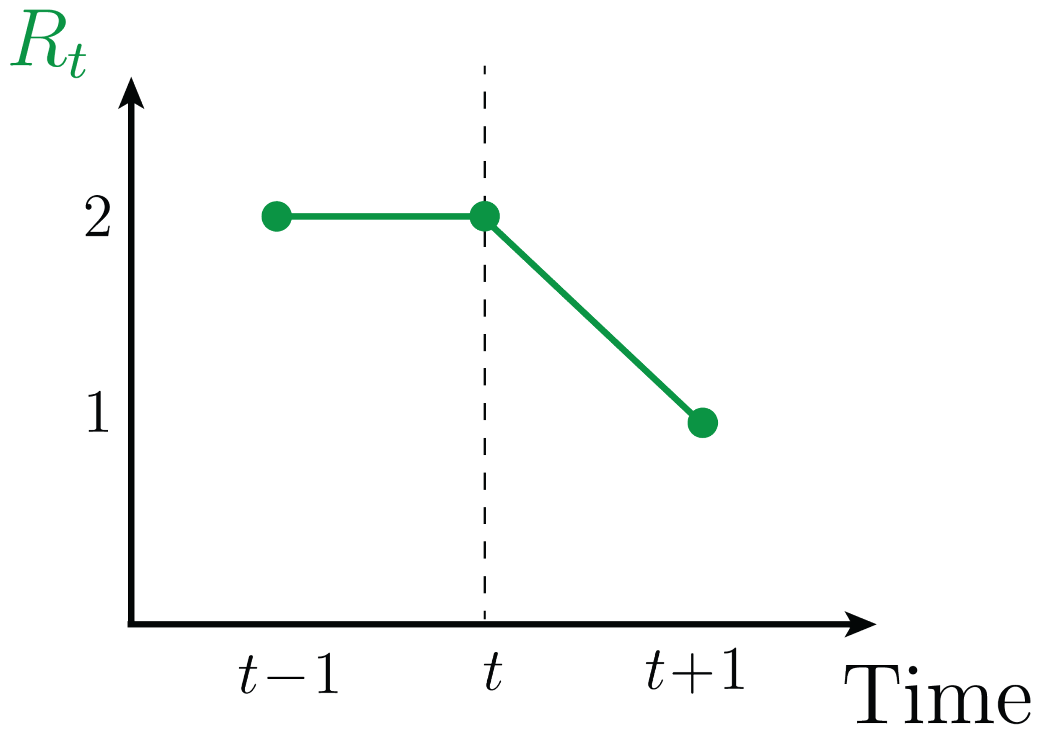


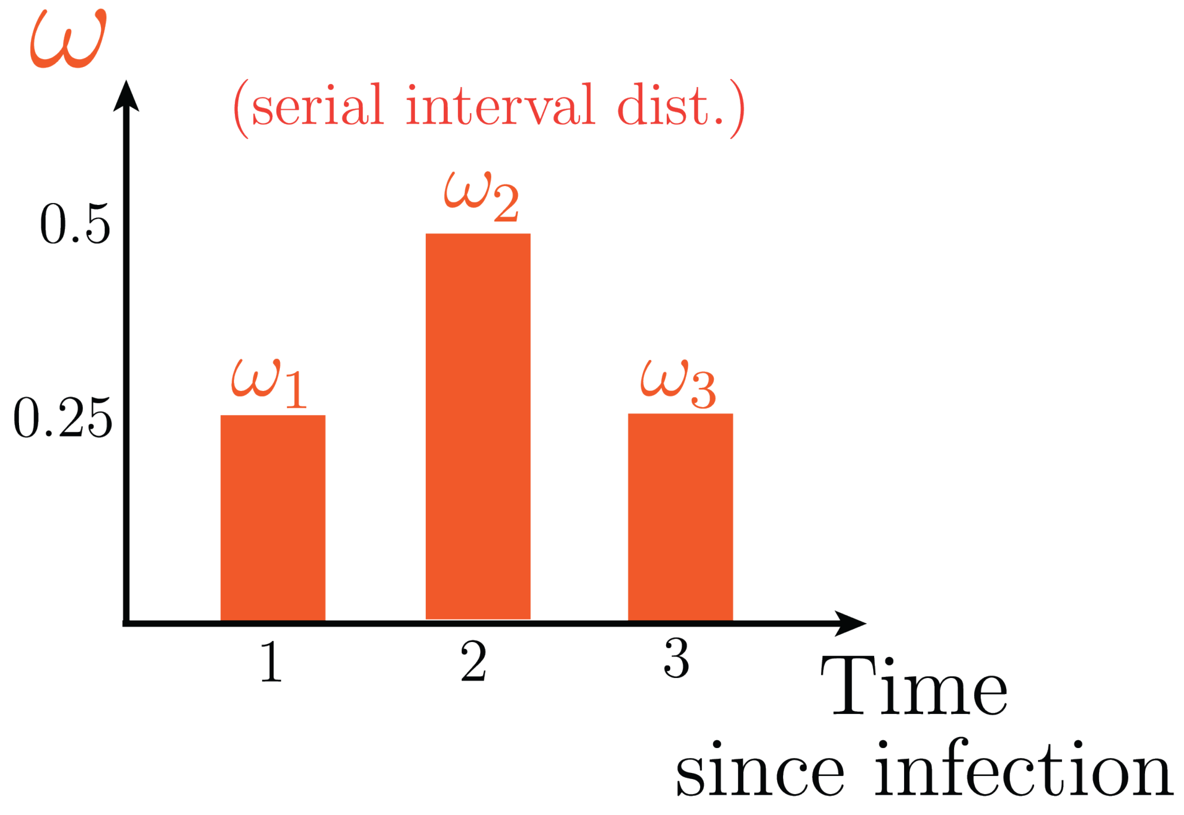


The renewal model

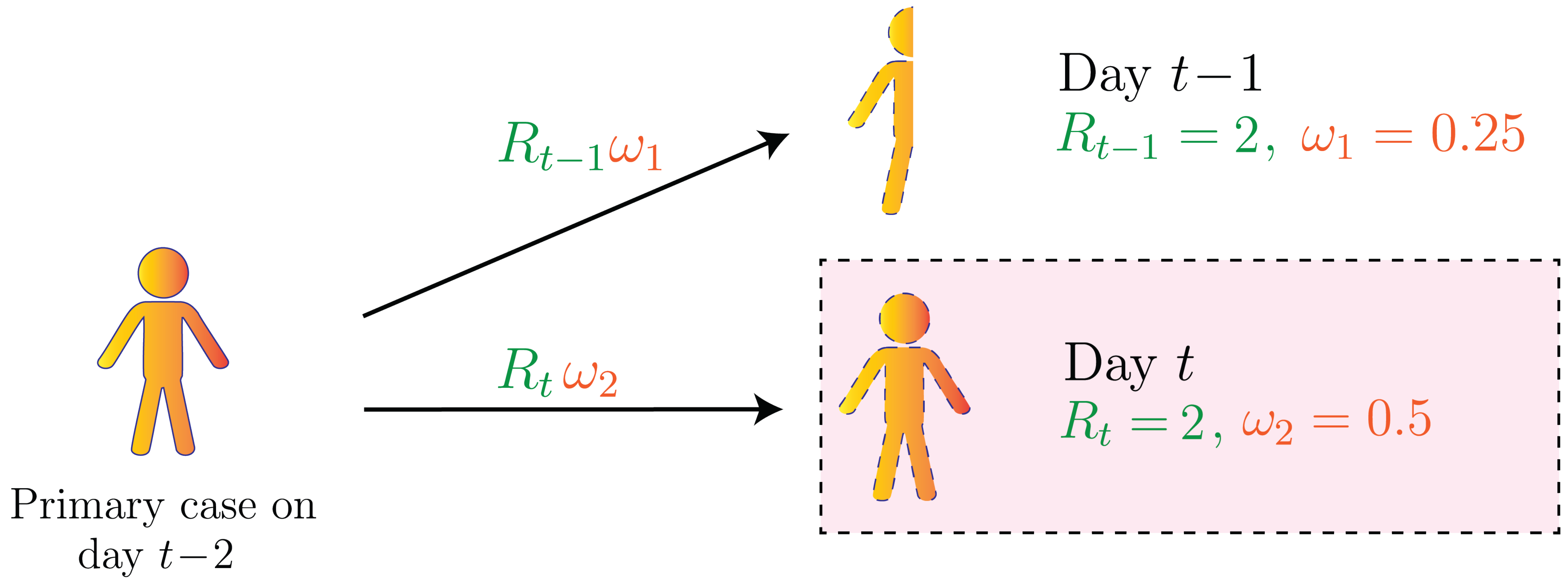


... will infect an average of 0.5 people yesterday, ...

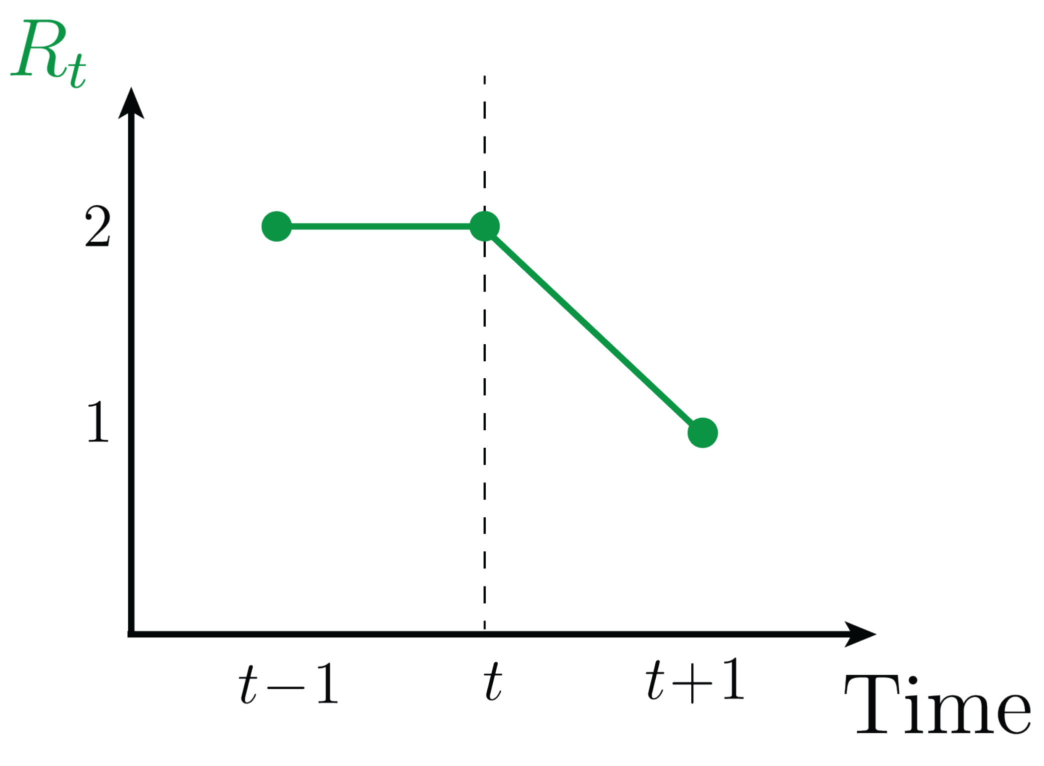


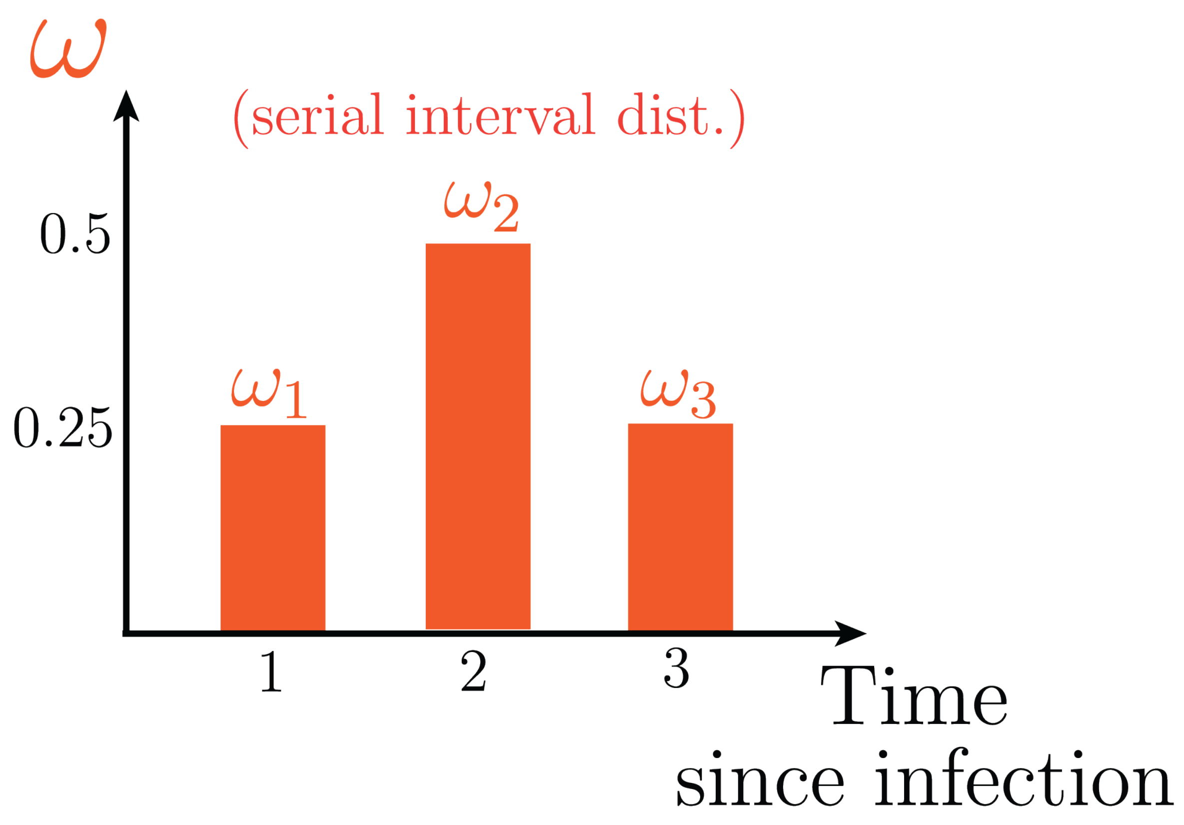


The renewal model

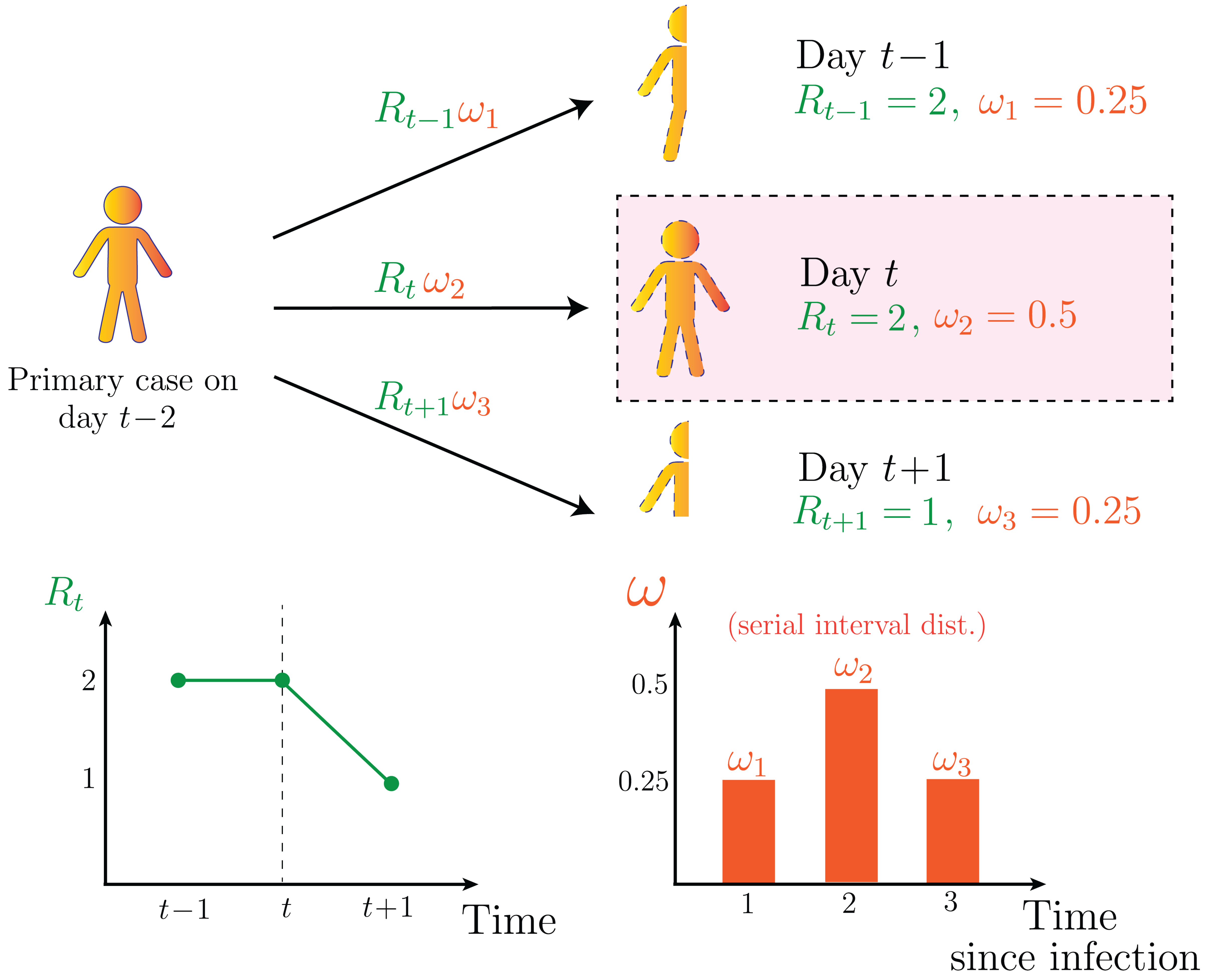


… an average of 1 person today, and…



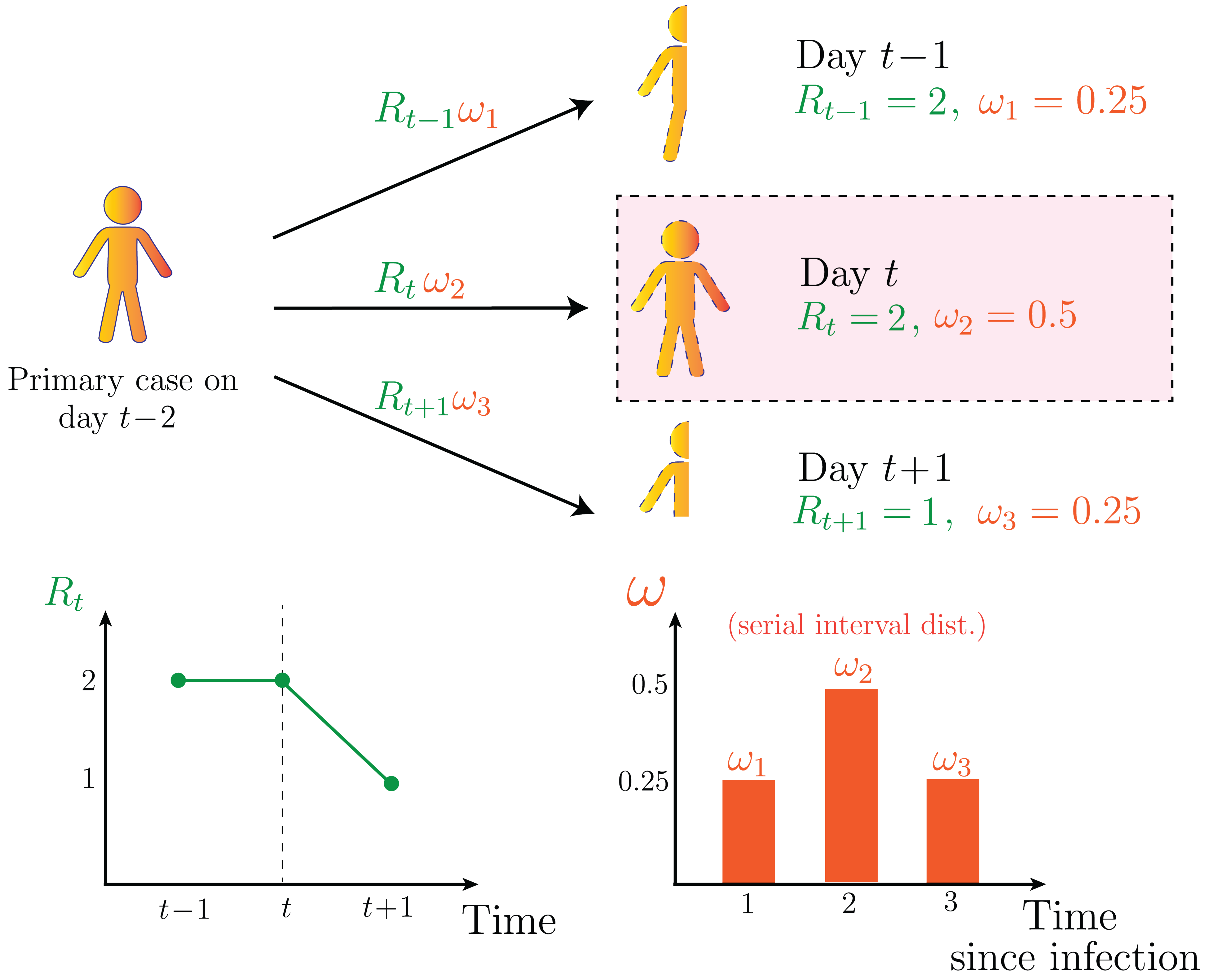


The renewal model



… an average of 0.25 people tomorrow.

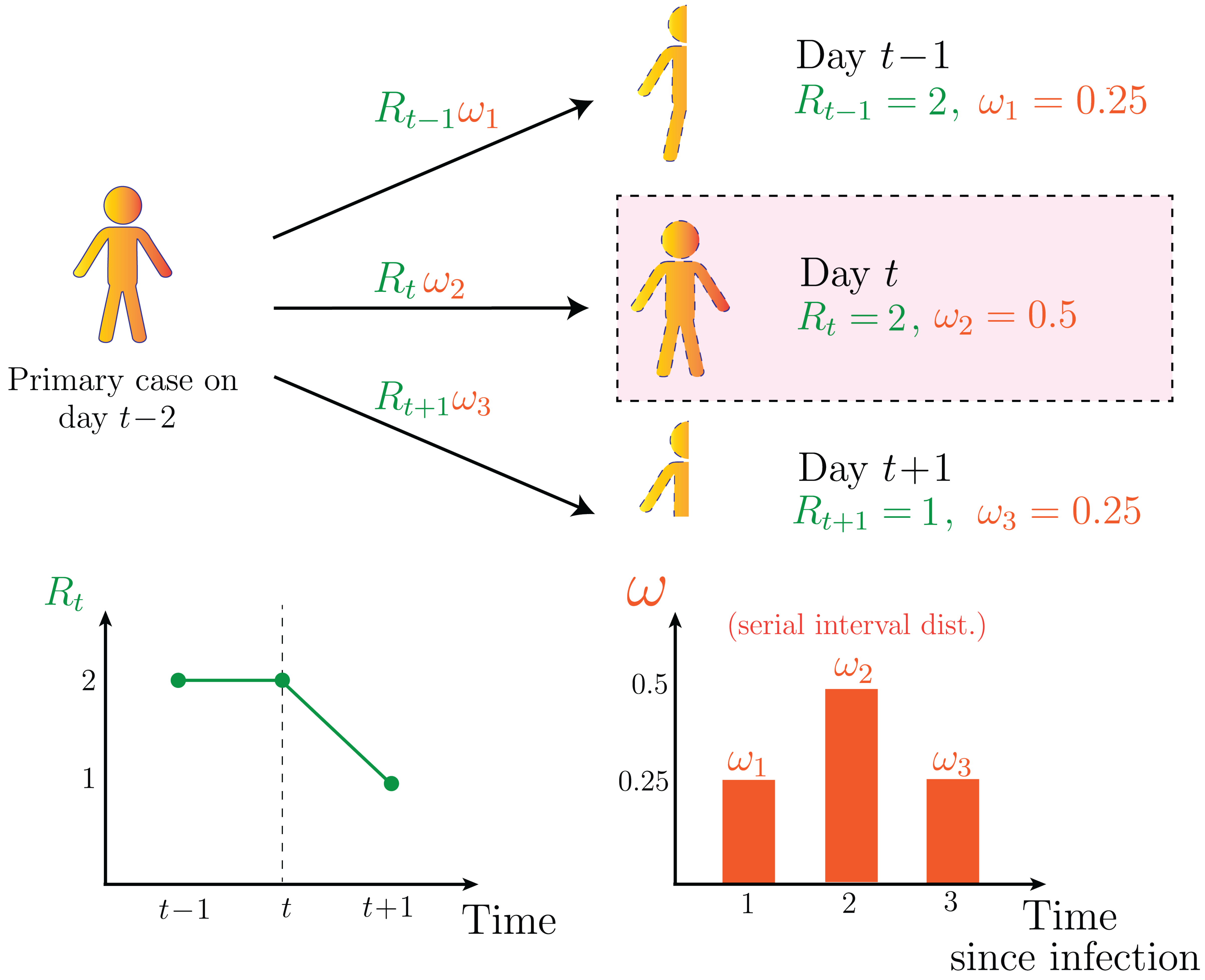
The renewal model

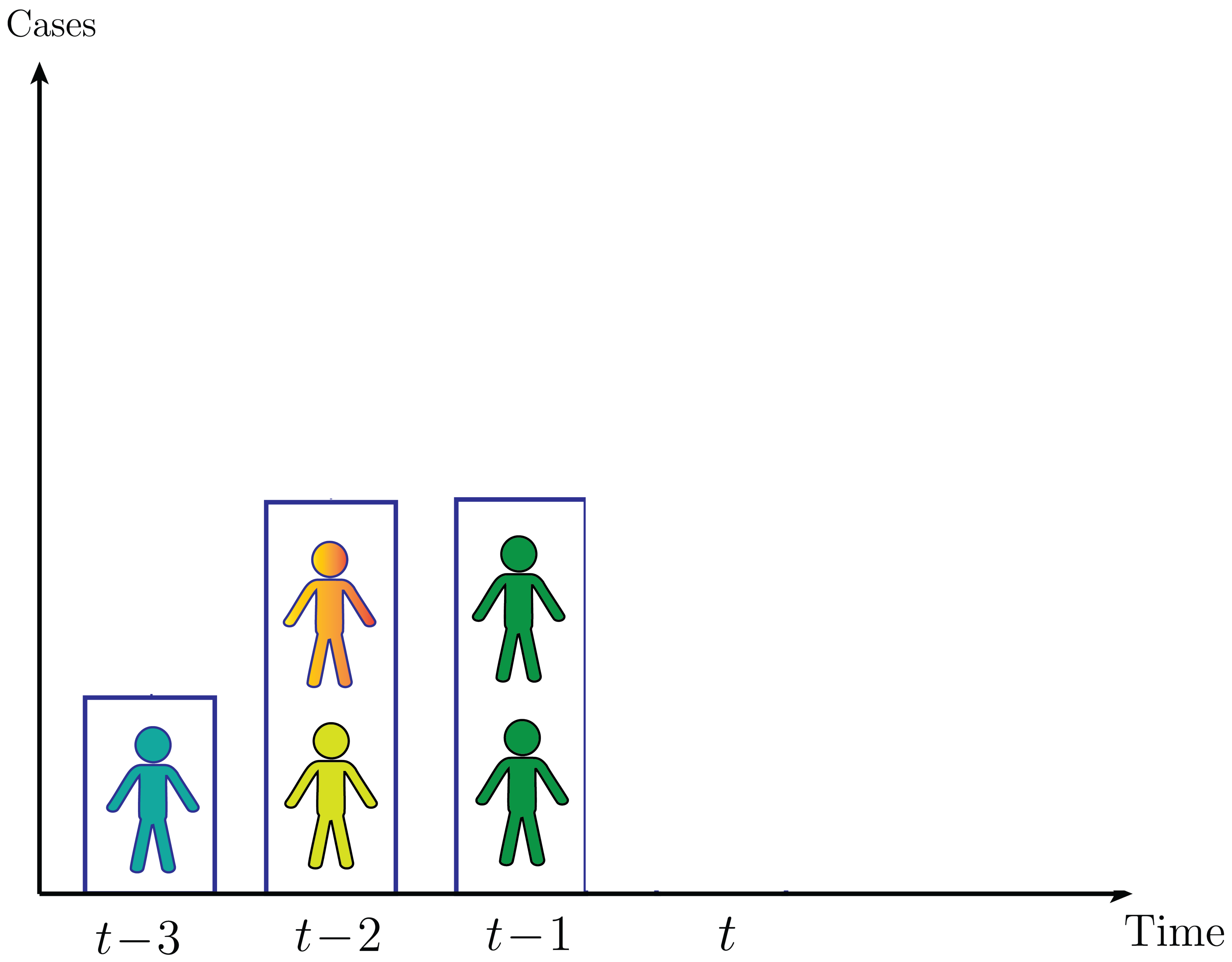


So...

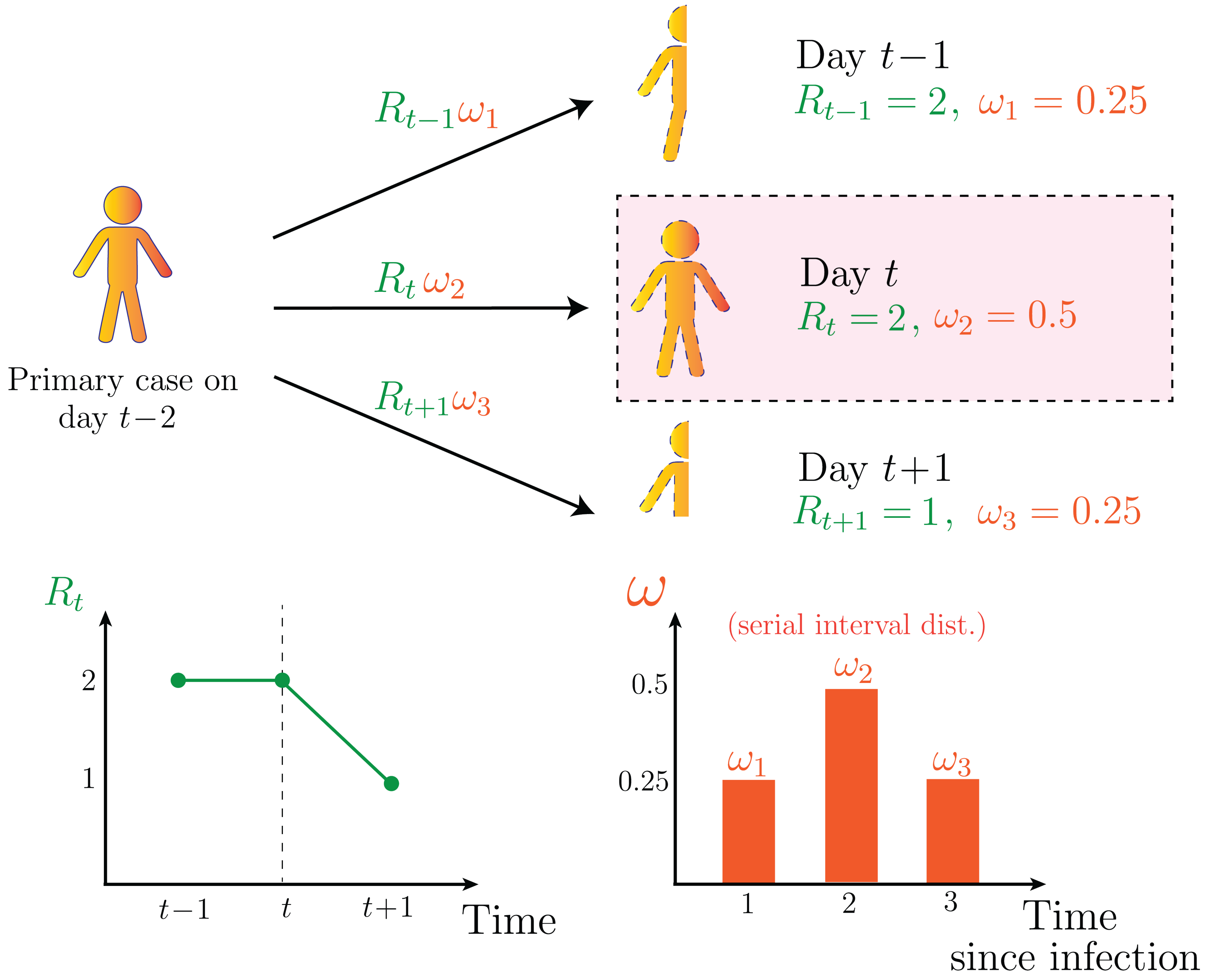
How many *total* cases do we expect today?

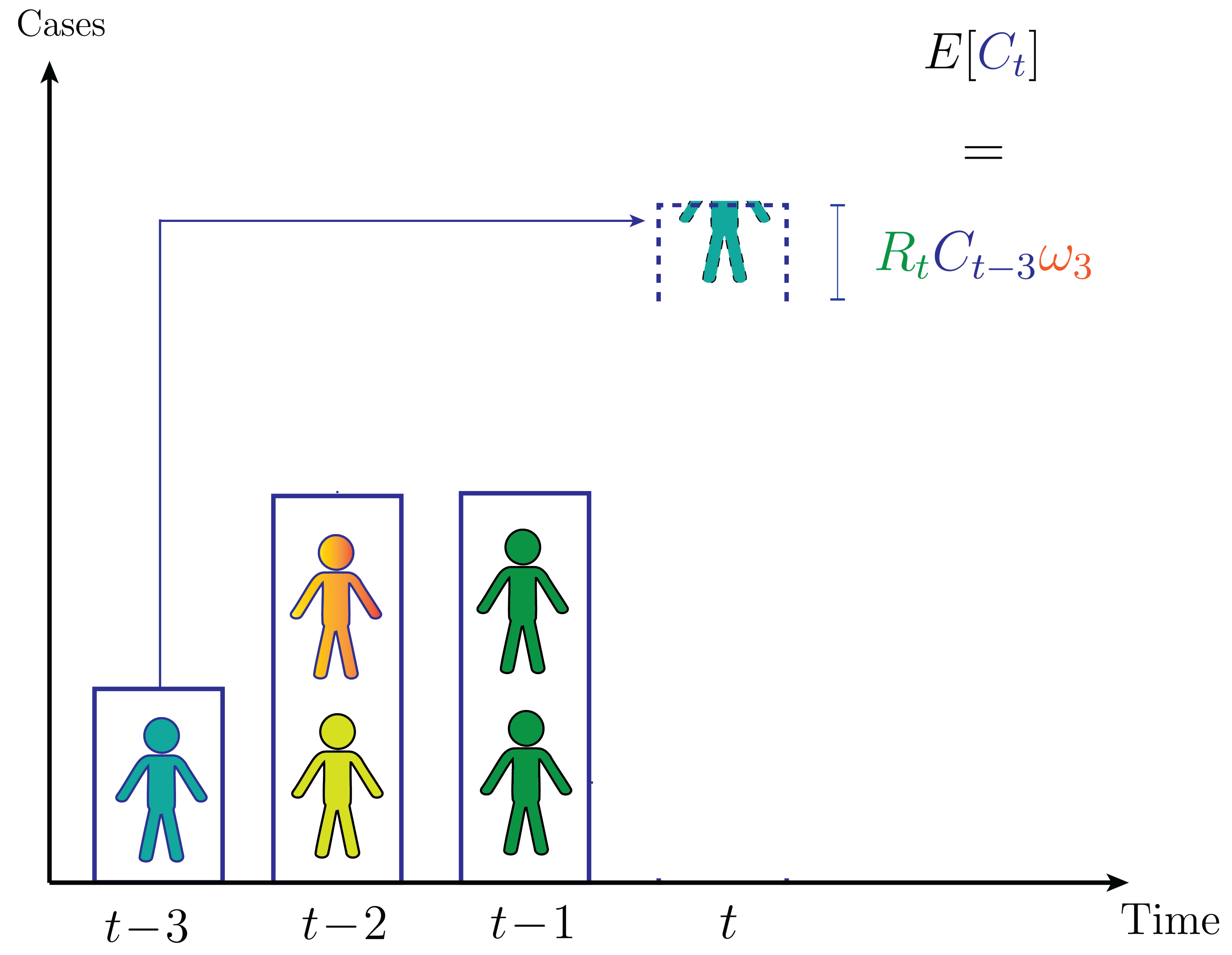
The renewal model



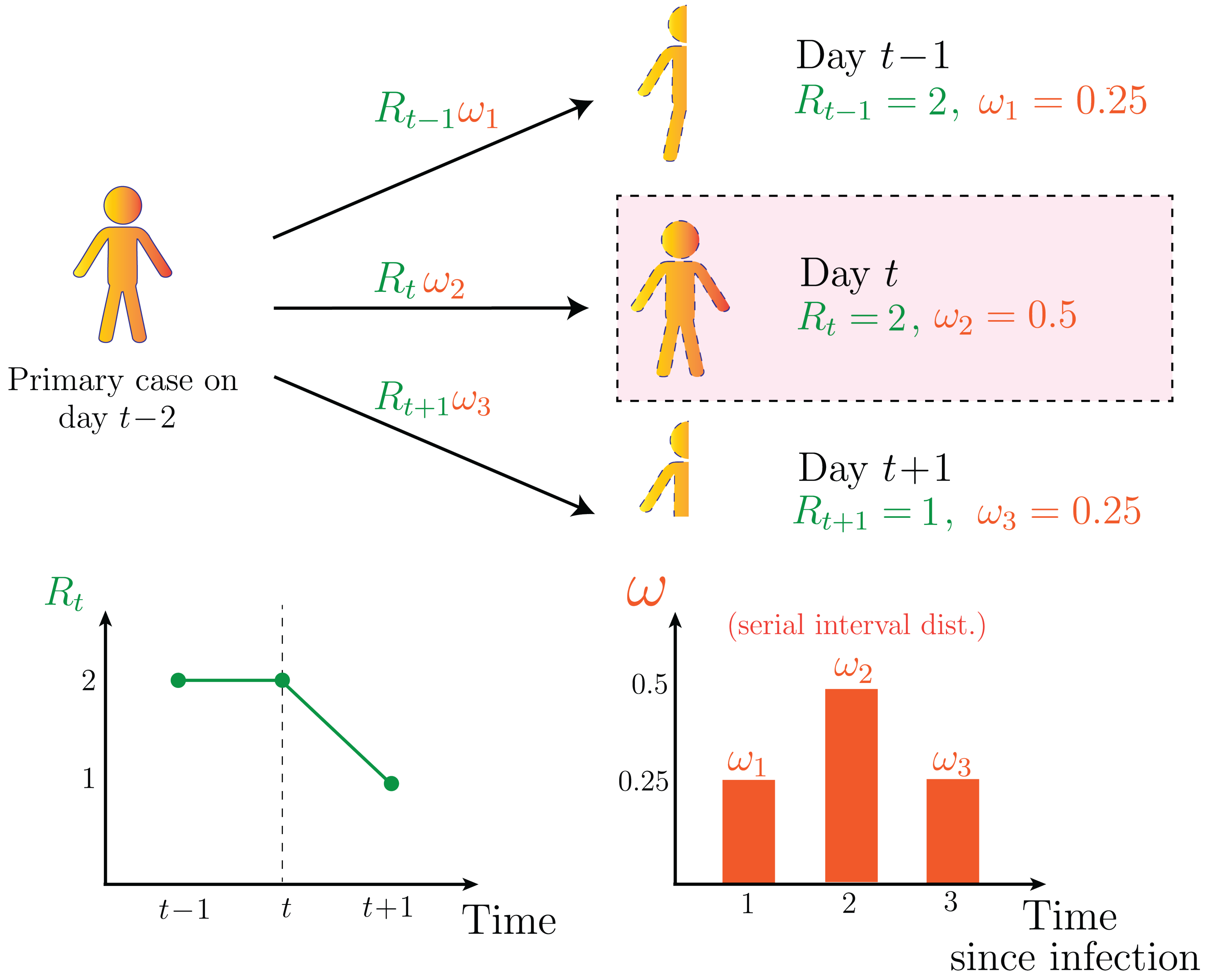


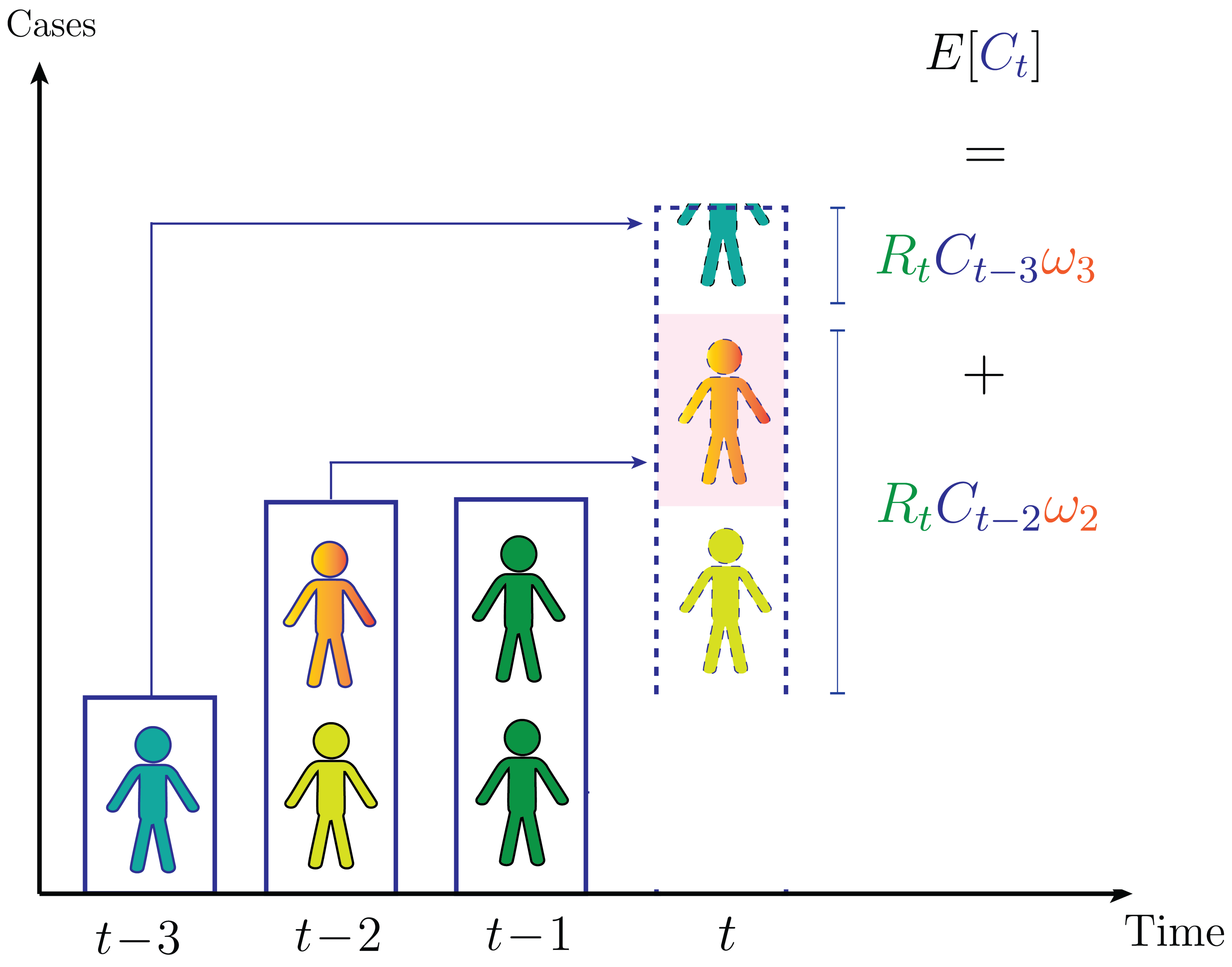
The renewal model



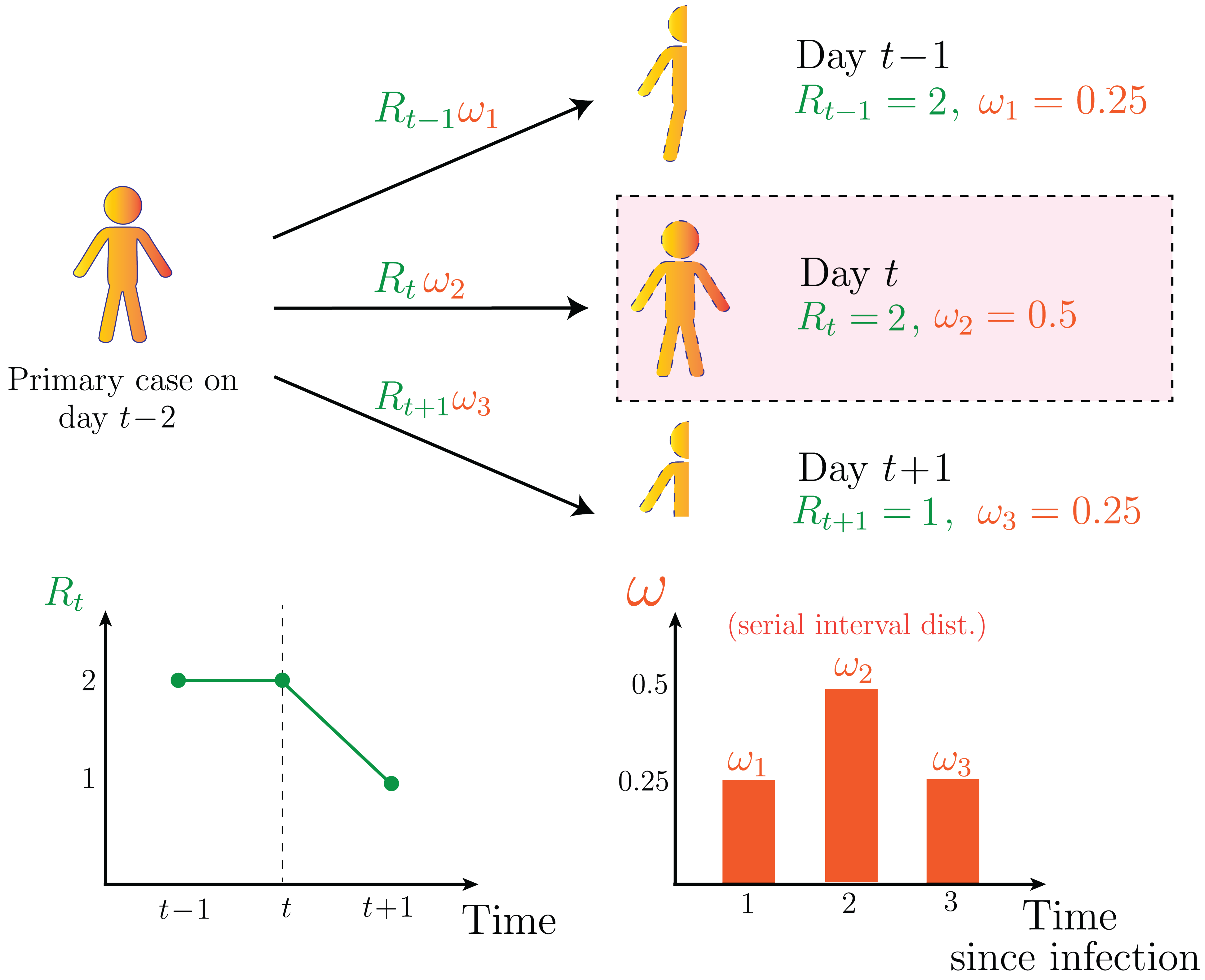


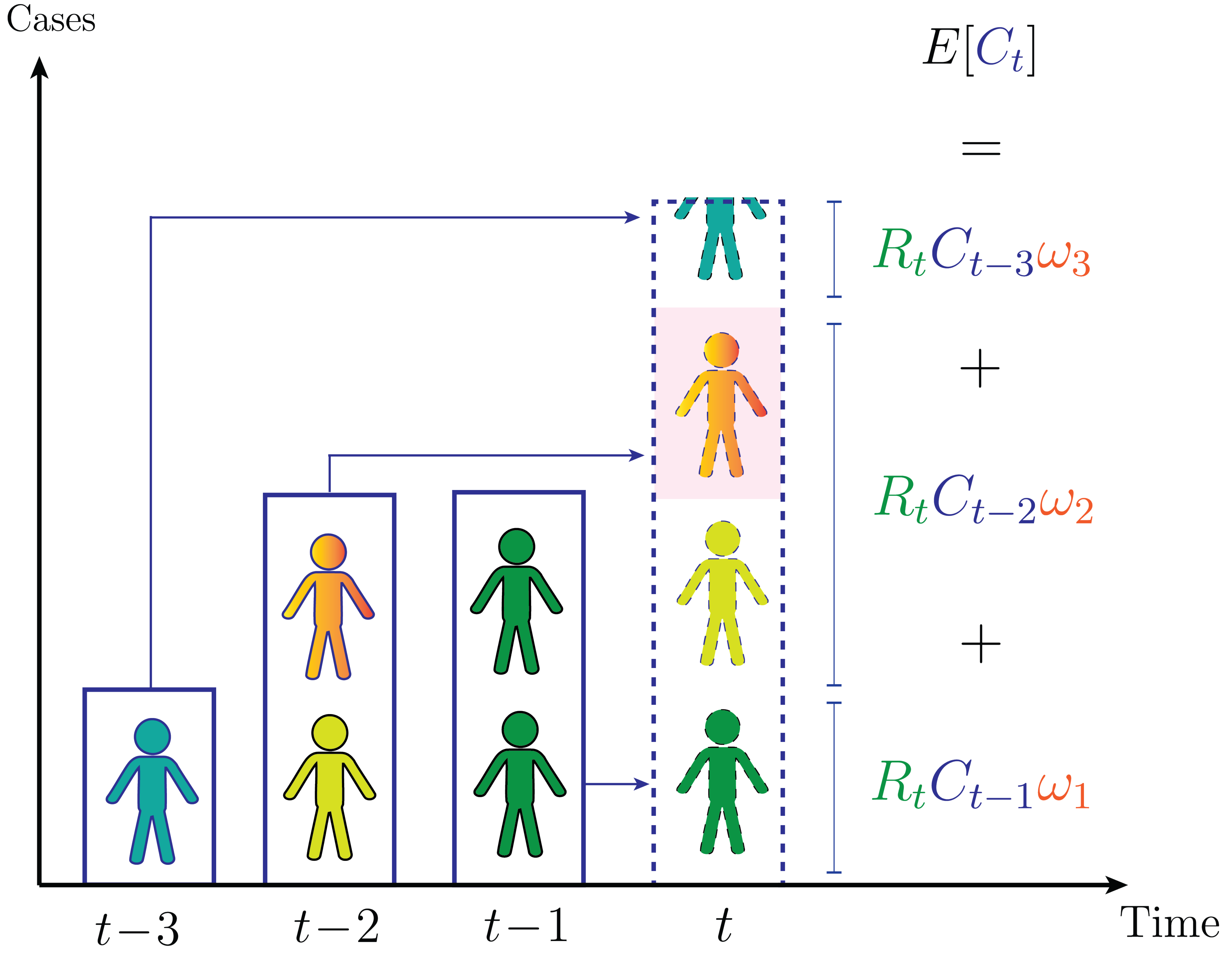
The renewal model





The renewal model





The renewal model

- This is a **very** flexible model

Prior for Rt

$$P(R_t)$$

$$\times$$

Likelihood

$$P(C_t \mid R_t, C_{1:t-1}) = Poiss\left(R_t \sum C_{t-u}\omega_u\right)$$

$$=$$

Posterior for Rt

$$P(R_t \mid C_{1:t})$$

Pointwise estimates of Rt

Dynamic model for Rt

$$P(R_t \mid R_{t-1}) = N(R_{t-1}, \sigma)$$

$$\times$$

Observation model

$$P(C_t \mid R_t, C_{1:t-1}) = Poiss\left(R_t \sum C_{t-u}\omega_u\right)$$

$$=$$

Posterior for Rt

$$P(R_t \mid C_{1:t})$$

Smoothed estimates of Rt

Dynamic model for Rt

$$P(R_t \mid R_{t-1}) = N(R_{t-1}, \sigma)$$

$$\times$$

Dynamic model for It

$$P(I_t \mid R_t, I_{1:t-1}) = Poiss\left(R_t \sum I_{t-u}\omega_u\right)$$

$$\times$$

Observation model

$$P(C_t \mid I_t) = Binomial\left(I_t, \rho\right)$$

$$=$$

Posterior for Rt

$$P(R_t \mid C_{1:t})$$

+ observation noise

Renewal Models

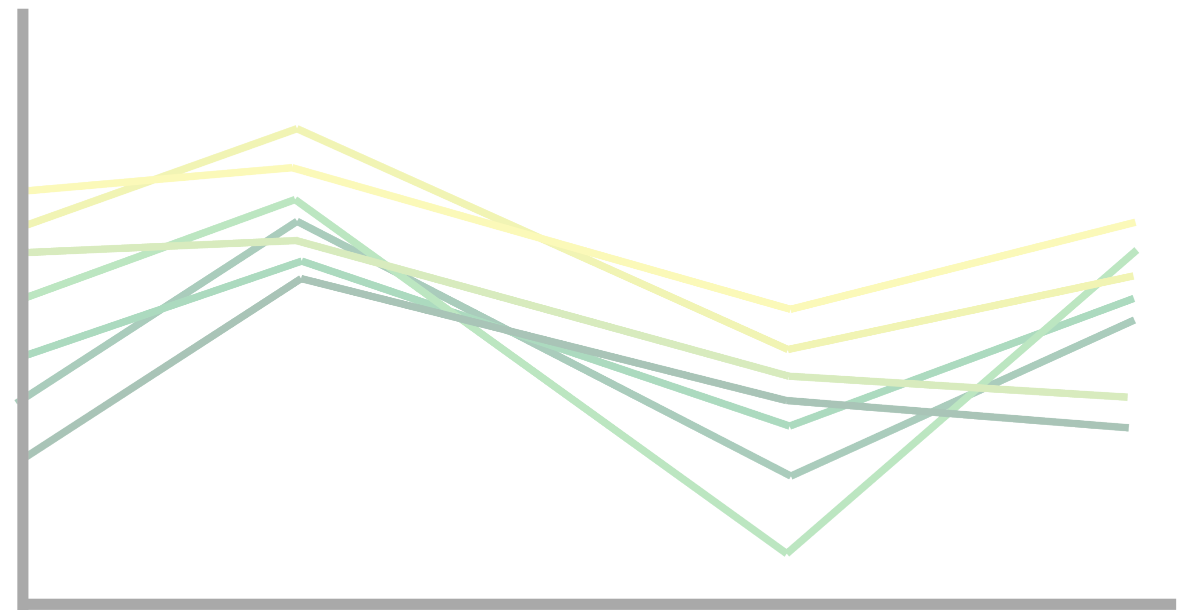
A simple ***model*** of an epidemic

$$C_t \sim \text{Poisson}\left(R_t \sum_{u=1}^{\omega_{max}} C_{t-u}\omega_u\right)$$

Used for reproduction number estimation

Also forecasting, elimination, effect of NPIs, etc

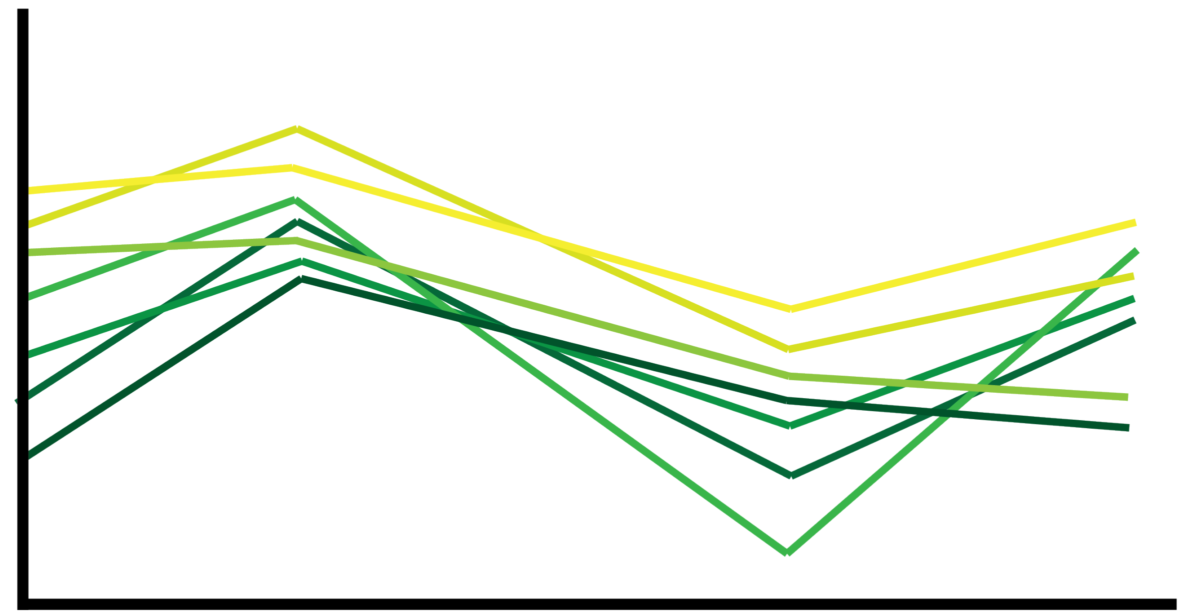
EpiEstim/EpiNow2/EpiFilter all use it



$$C_t \sim \text{Poisson}\left(R_t \sum_{u=1}^{\omega_{max}} C_{t-u}\omega_u\right)$$

Sequential Monte Carlo

A ***method*** for fitting hidden-state models



Also known as “particle filters”

Very flexible!

Can account for many biases at once

Sequential Monte Carlo

- Origins in **importance sampling** and **Bayesian filtering/smoothing** (e.g. Kalman filters)
- **Bootstrap filter (Gordon, 1993)**: the first method for non-linear non-Gaussian state-space models
- Then advanced by Doucet, Del Moral, Chopin, Kantas, Andrieu and many others*
- We consider **hidden-state estimation** and **parameter estimation** separately

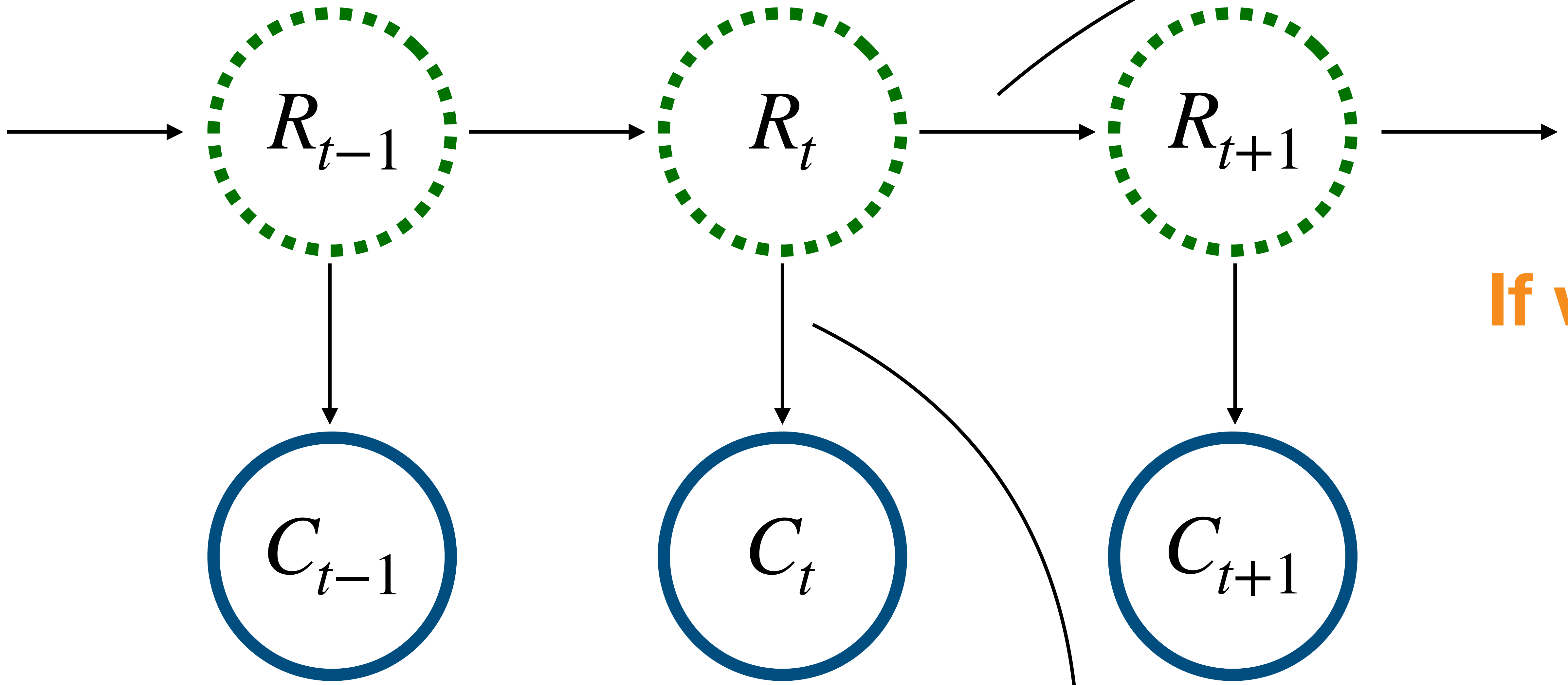
(These names are listed as they are associated with helpful tutorial/summary/overview papers!)

Sequential Monte Carlo: the bootstrap filter…

…fits a **hidden-state model**

Dynamic model for Rt

$P(R_t \mid R_{t-1})$



If we can sample from this

and evaluate this,

Observation model for cases

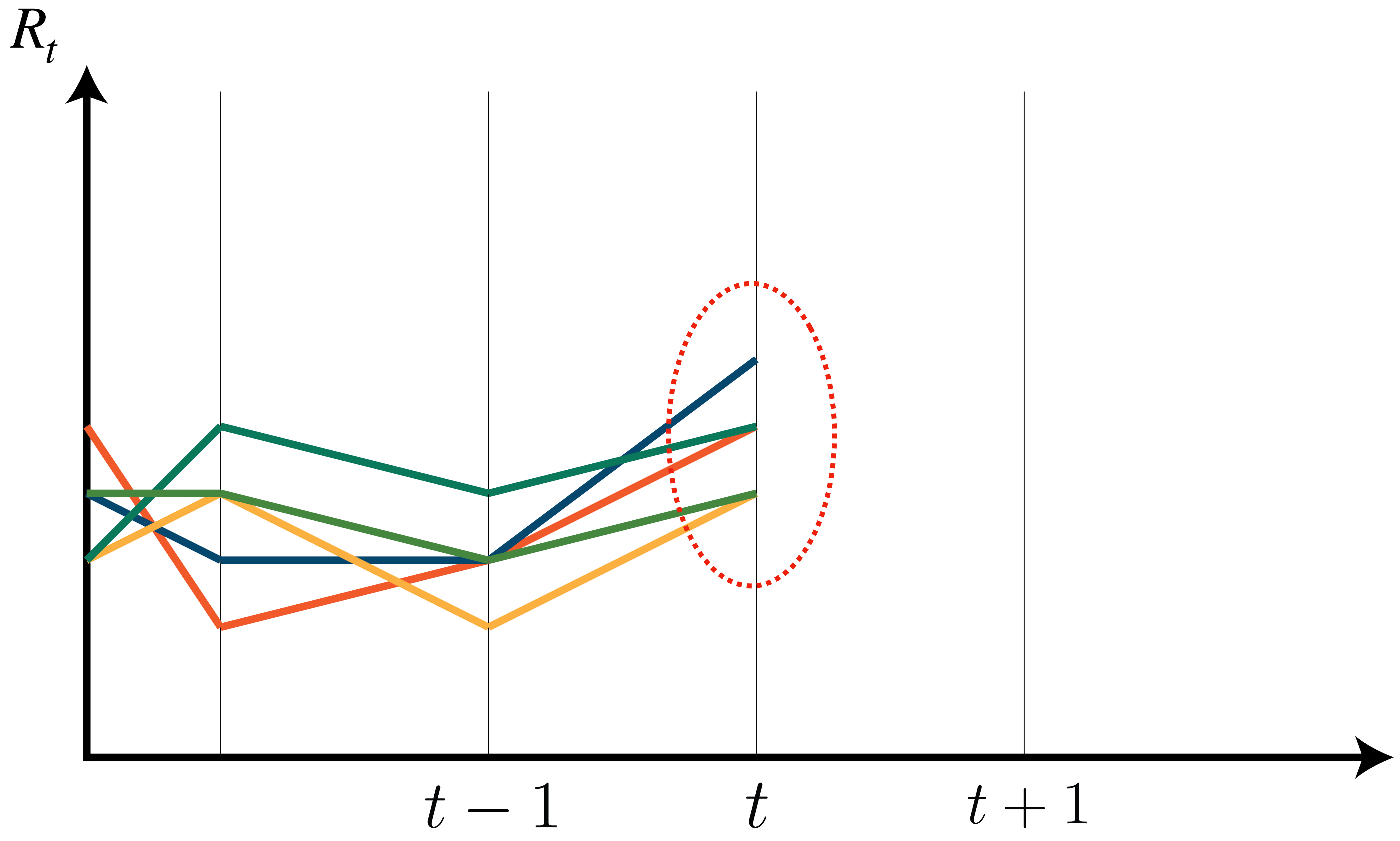
$P(C_t \mid R_t, C_{1:t-1})$

then we can use SMC methods.

Sequential Monte Carlo: the bootstrap filter

Assume we have samples:

$$R_t^{(i)} \sim P(R_t \mid C_{1:t}), \quad i = 1,\dots,N$$



We update these samples at $t+1$ by:

1. **Projecting** according to the dynamic model

$$\tilde{R}_{t+1}^{(i)} \sim P(R_{t+1} \mid R_t^{(i)})$$

2. **Weighting** according to the observation model

$$W_{i,t+1} = P(C_{t+1} \mid \tilde{R}_{t+1}^{(i)}, C_{1:t})$$

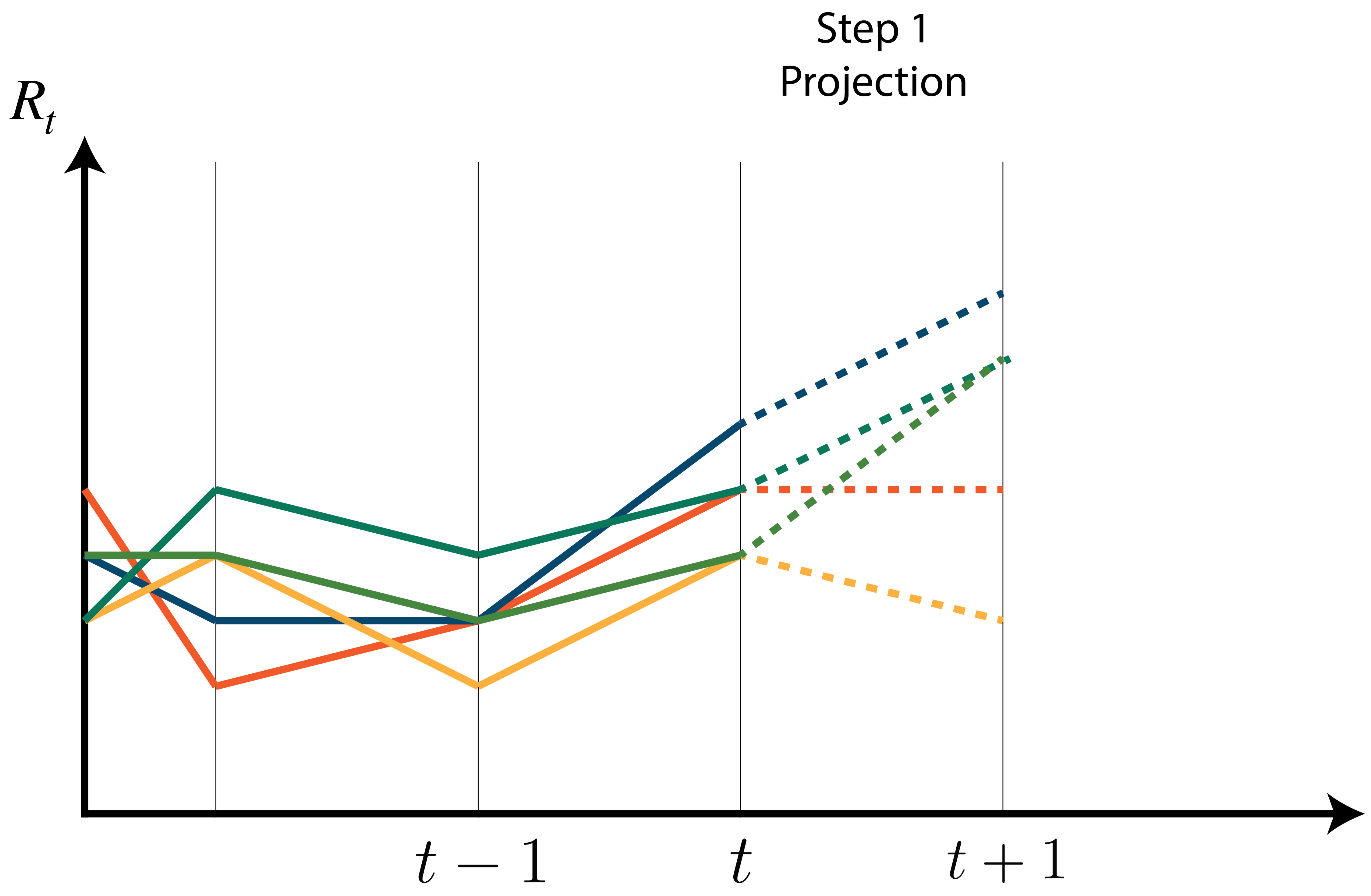
3. **Resampling** projected “particles”, giving:

$$R_{t+1}^{(i)} \sim P(R_{t+1} \mid C_{1:t+1}), \quad i = 1,\dots,N$$

Sequential Monte Carlo: the bootstrap filter

Assume we have samples:

$$R_t^{(i)} \sim P(R_t \mid C_{1:t}), \quad i = 1,\ldots,N$$



We update these samples at $t+1$ by:

1. **Projecting** according to the dynamic model

$$\tilde{R}_{t+1}^{(i)} \sim P(R_{t+1} \mid R_t^{(i)})$$

2. **Weighting** according to the observation model

$$W_{i,t+1} = P(C_{t+1} \mid \tilde{R}_{t+1}^{(i)}, C_{1:t})$$

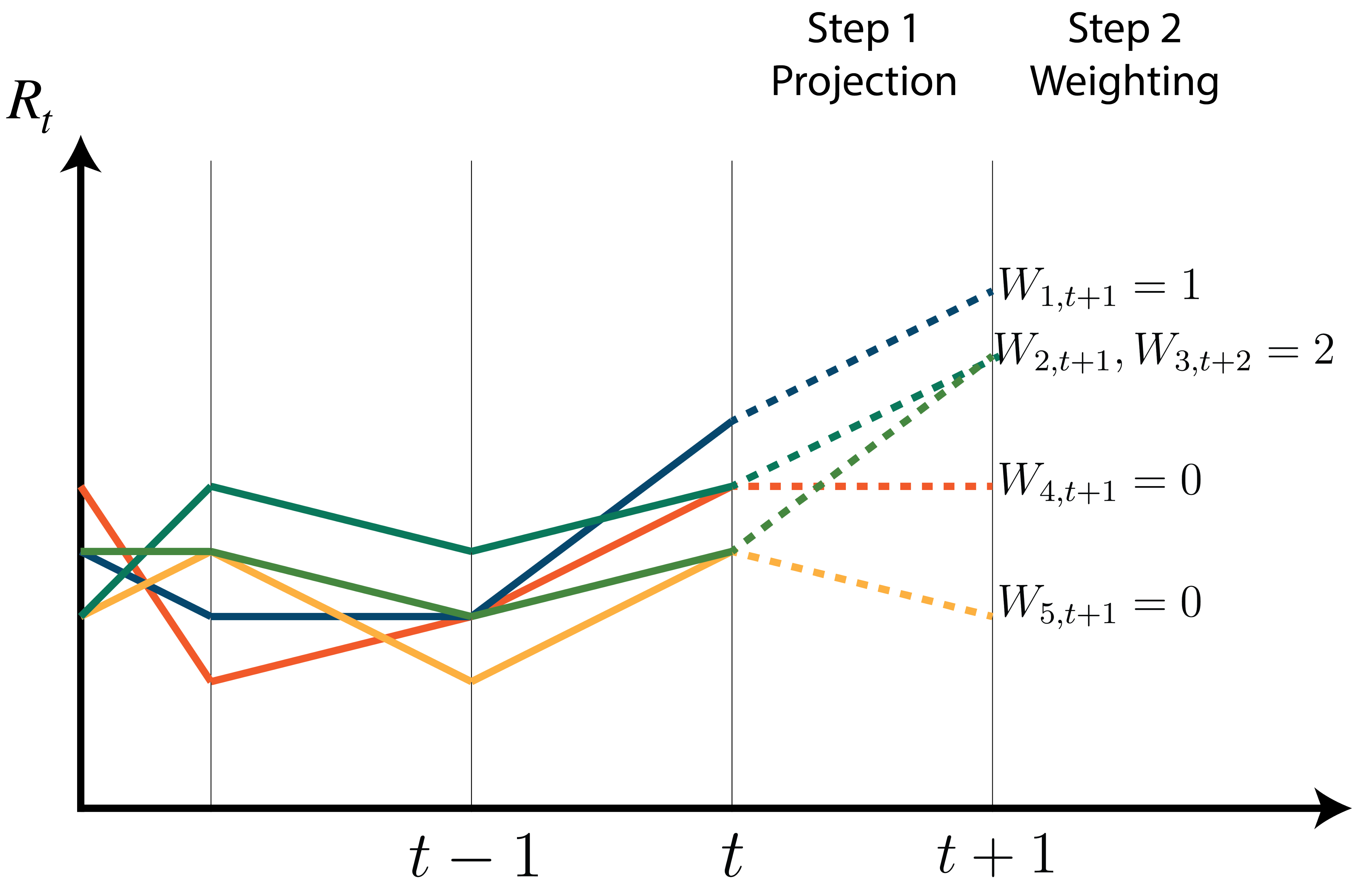
3. **Resampling** projected "particles", giving:

$$R_{t+1}^{(i)} \sim P(R_{t+1} \mid C_{1:t+1}), \quad i = 1,\ldots,N$$

Sequential Monte Carlo: the bootstrap filter

Assume we have samples:

$$R_t^{(i)} \sim P(R_t \mid C_{1:t}), \quad i = 1,\ldots,N$$



We update these samples at $t+1$ by:

1. **Projecting** according to the dynamic model

$$\tilde{R}_{t+1}^{(i)} \sim P(R_{t+1} \mid R_t^{(i)})$$

2. **Weighting** according to the observation model

$$W_{i,t+1} = P(C_{t+1} \mid \tilde{R}_{t+1}^{(i)}, C_{1:t})$$

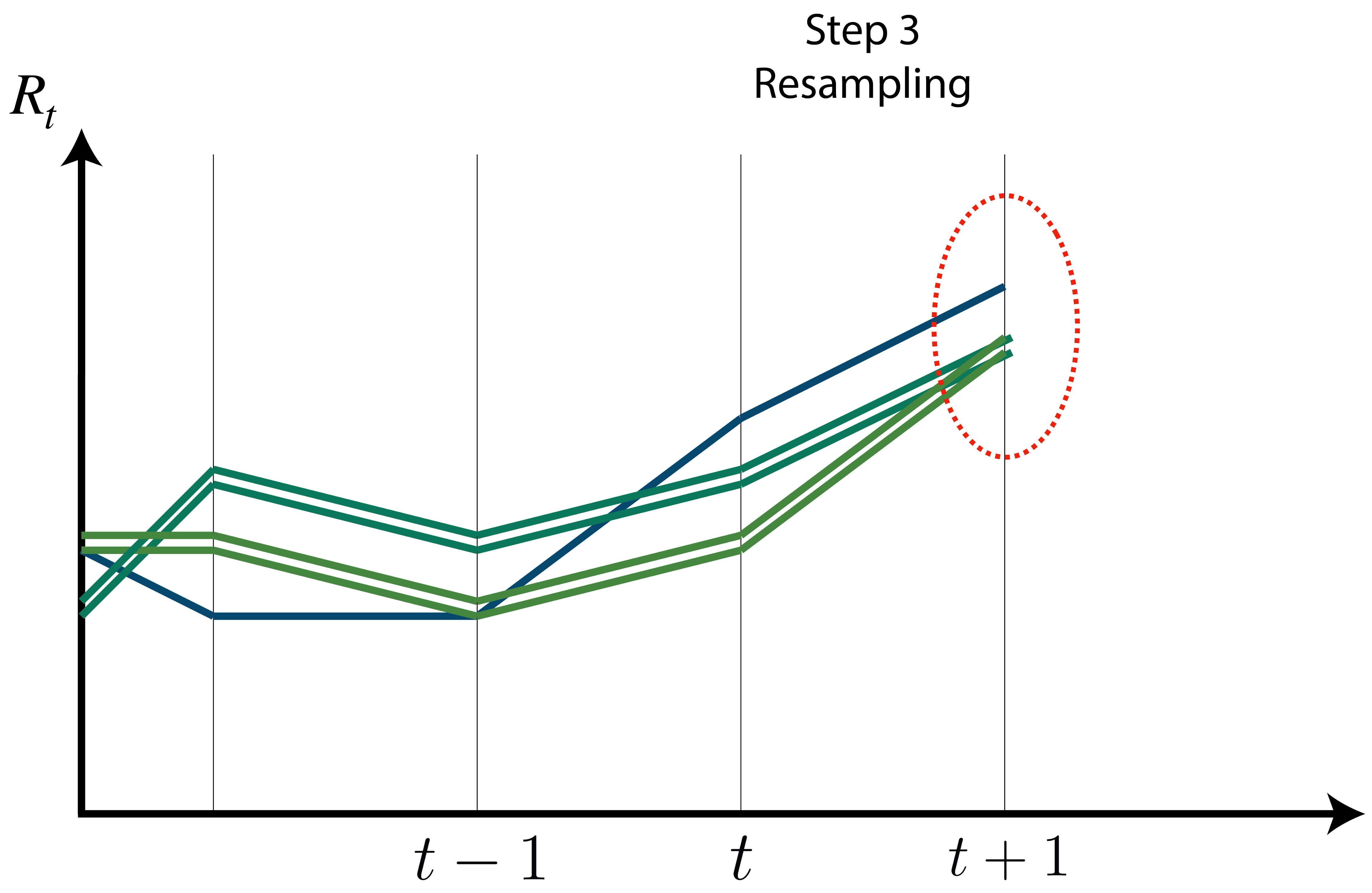
3.**Resampling** projected "particles", giving:

$$R_{t+1}^{(i)} \sim P(R_{t+1} \mid C_{1:t+1}), \quad i = 1,\ldots,N$$

Sequential Monte Carlo: the bootstrap filter

Assume we have samples:

$$R_t^{(i)} \sim P(R_t \,|\, C_{1:t}), \quad i = 1,\ldots,N$$



We update these samples at $t+1$ by:

1. **Projecting** according to the dynamic model

$$\tilde{R}_{t+1}^{(i)} \sim P(R_{t+1} \,|\, R_t^{(i)})$$

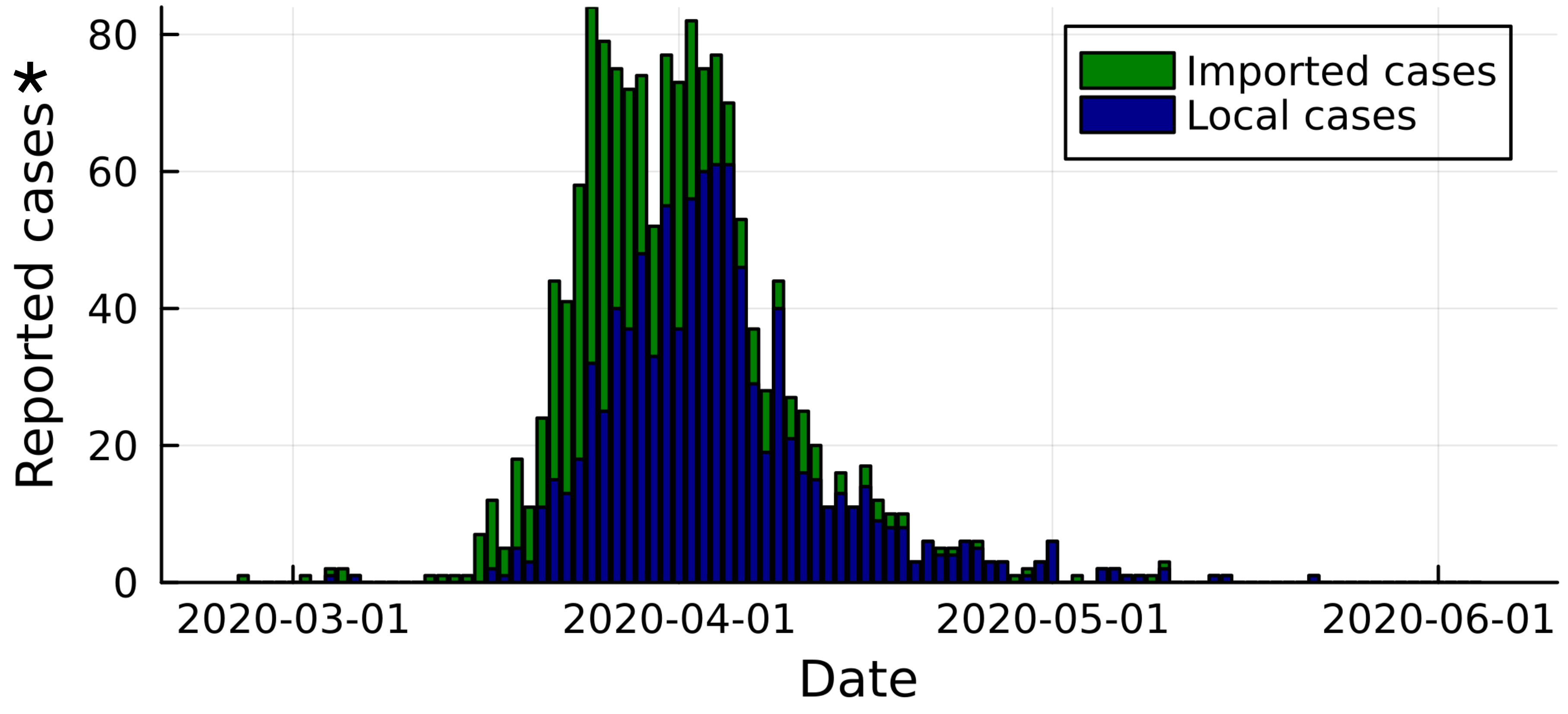
2. **Weighting** according to the observation model

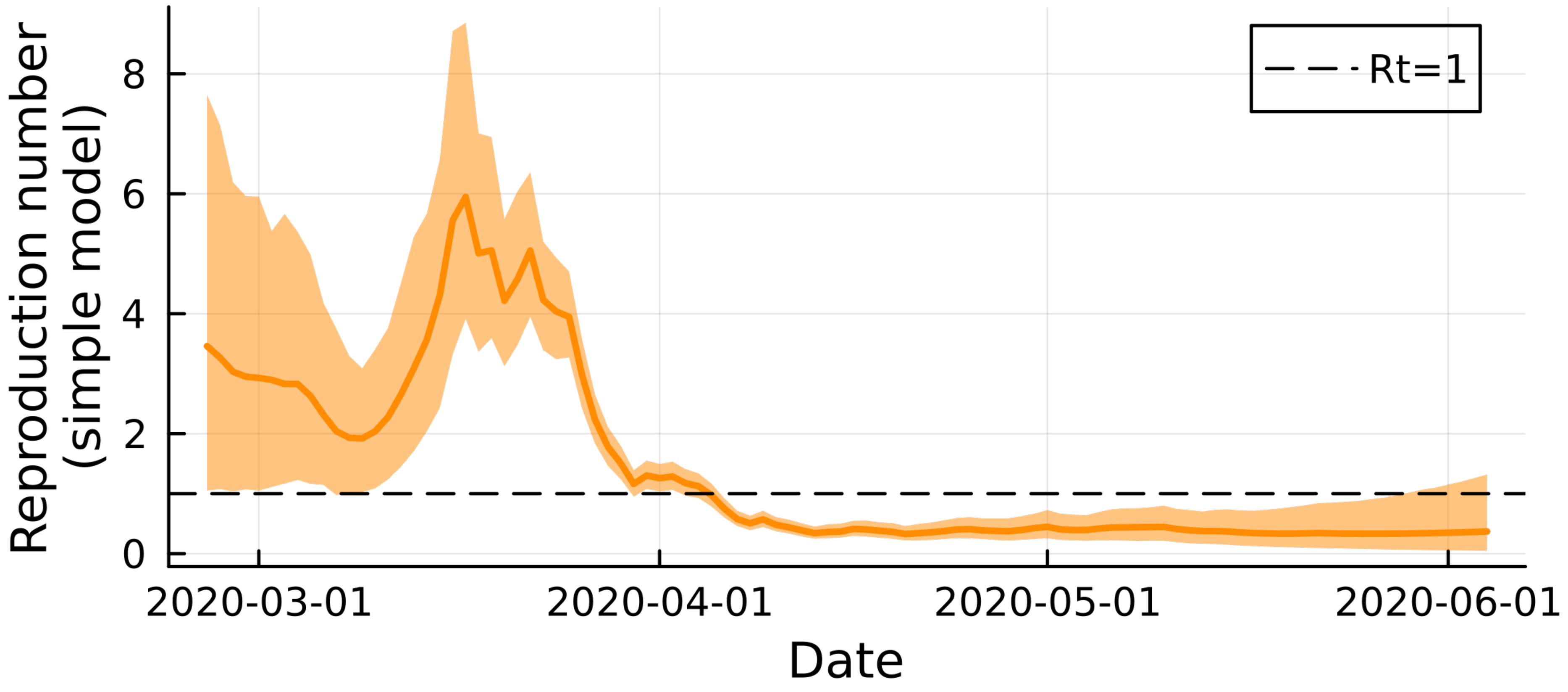
$$W_{i,t+1} = P(C_{t+1} \,|\, \tilde{R}_{t+1}^{(i)}, C_{1:t})$$

3.**Resampling** projected "particles", giving:

$$R_{t+1}^{(i)} \sim P(R_{t+1} \,|\, C_{1:t+1}), \quad i = 1,\ldots,N$$

A simple example…





** Reported case data are from the first 100 days of the COVID-19 outbreak in New Zealand*

Dynamic model for R_t

$$\log R_t \,|\, R_{t-1} \sim \text{Normal}(\log R_{t-1}, \sigma)$$

Observation model

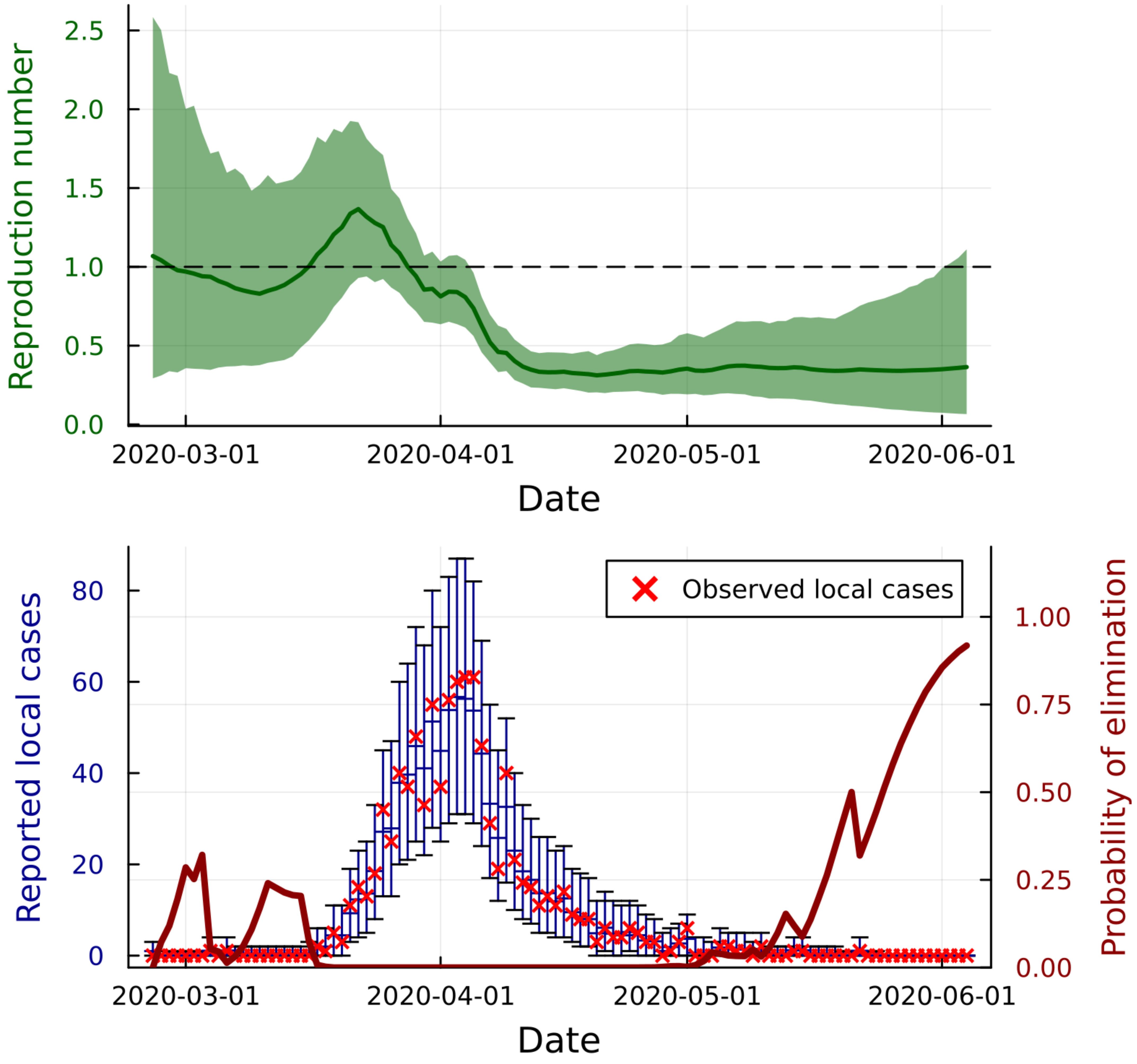
$$C_t \sim \text{Poisson}\left(R_t \sum_{u=1}^{\omega_{max}} C_{t-u}\omega_u \right)$$

The entire algorithm requires only 11 lines of code:

```
# Setup:
N = 100000 # Number of particles
ω = pdf.(Gamma(2.36, 2.74), 1:100) # Serial interval
Y = loadData("NZCOVID") # Load data
R = zeros(N, 100) # Matrix to store particle values (of log Rt)

# Run the bootstrap filter:
R[:,1] = log.(rand(Uniform(0, 10), N)) # Sample initial values
for tt = 2:100
    R[:,tt] = rand.(Normal.(R[:,tt-1], 0.2)) # Project (log) Rt
    Λ = sum(Y.Ct[(tt-1):-1:1] .* ω[1:(tt-1)]) # Calculate the force of infection
    W = pdf.(Poisson.(Λ * exp.(R[:,tt])), Y.Ct[tt]) # Calculate weights
    R[1:N, max(tt-40, 1):tt] = R[wsample(1:N, W, N), max(tt-40, 1):tt] # Resample
end
```

Imported cases, reporting noise, and elimination



Dynamic model for R_t

$$\log R_t \mid R_{t-1} \sim \text{Normal}(\log R_{t-1}, \sigma)$$

$$I_t \mid R_t, I_{1:t-1} \sim \text{Poisson}\left(R_t \sum_{u=1}^{\omega_{max}} (I_{t-u} + M_{t-u})\omega_u \right)$$

Observation model

$$C_t \sim \text{Negative binomial}(\cdots)$$

Probability of elimination estimated by simulating the model forward two weeks and checking whether any new infections occur.

Sequential Monte Carlo: Parameter estimation

- A much harder and more expensive problem!
- We use Particle Marginal Metropolis Hastings (PMMH)

Dynamic model for R_t

$$\log R_t \mid R_{t-1} \sim \text{Normal}(\log R_{t-1}, \sigma)$$

Observation model

$$C_t \sim \text{Poisson}\left(R_t \sum_{u=1}^{\omega_{max}} C_{t-u}\omega_u\right)$$

Sequential Monte Carlo: Parameter estimation

- Predictive decomposition of the likelihood:

$$L(\theta) = P(C_{1:T} \mid \theta) = P(C_1 \mid \theta) \prod_{t=2}^{T} \underbrace{P(C_t \mid C_{1:t-1}, \theta)}$$

- Where the one-step-ahead likelihood is just the average of the bootstrap filter weights:

$$\overbrace{P(C_t \mid C_{1:t-1}, \theta)} = E_{R_t \mid C_{1:t-1}}[P(C_t \mid R_t, C_{1:t-1}, \theta)] \approx \frac{1}{N} \sum_{i=1}^{N} W_{t,i}$$

Sequential Monte Carlo: Parameter estimation

- To estimate the log-likelihood, we run the bootstrap filter and calculate:

$$\hat{\ell}(\theta) = \sum_{t=1}^{T} \log \bar{W}_t$$

- Then just plug this into a standard Metropolis Hastings algorithm!

- Need to be careful about $S.D.(\hat{\ell}(\theta))$….

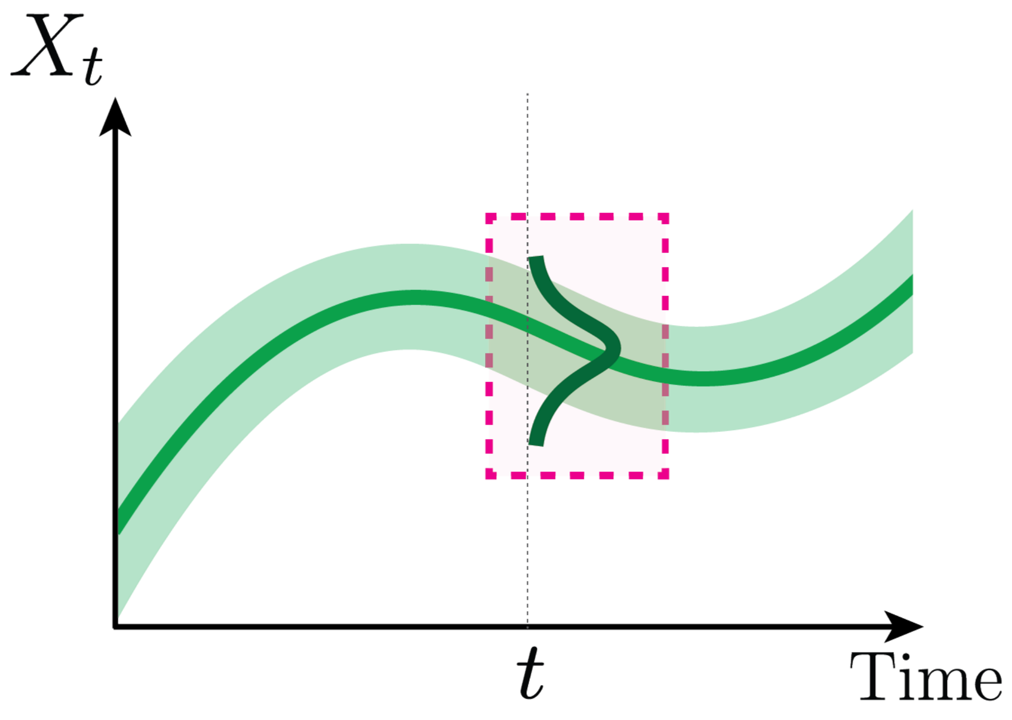
Sequential Monte Carlo: Overall approach

Bootstrap filter

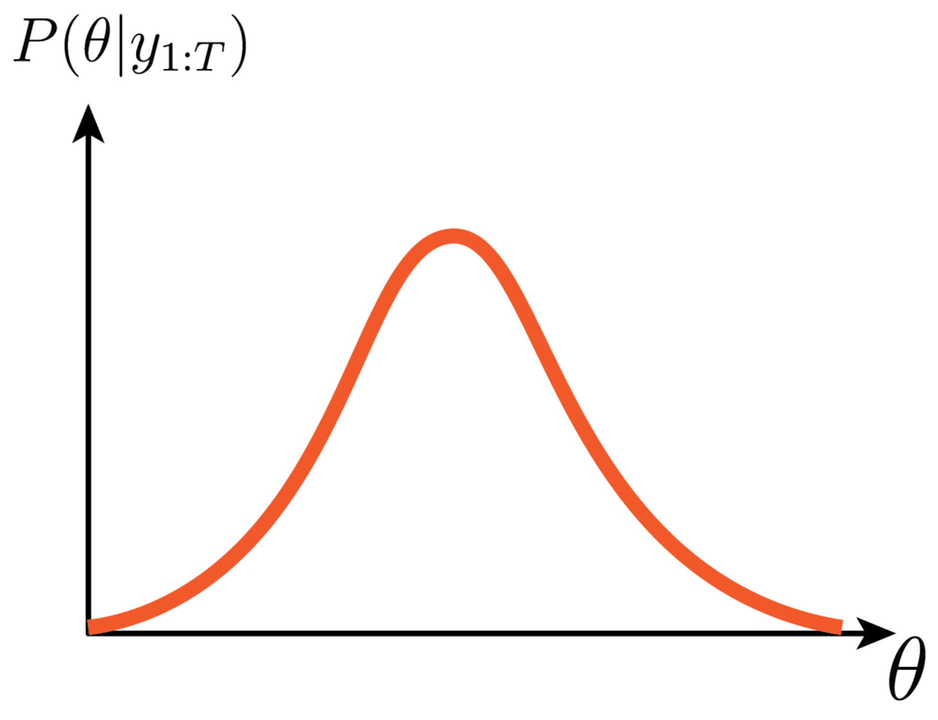
Posterior distribution of hidden states

$P(X_t|y_{1:T}, \theta)$

(conditional on θ)

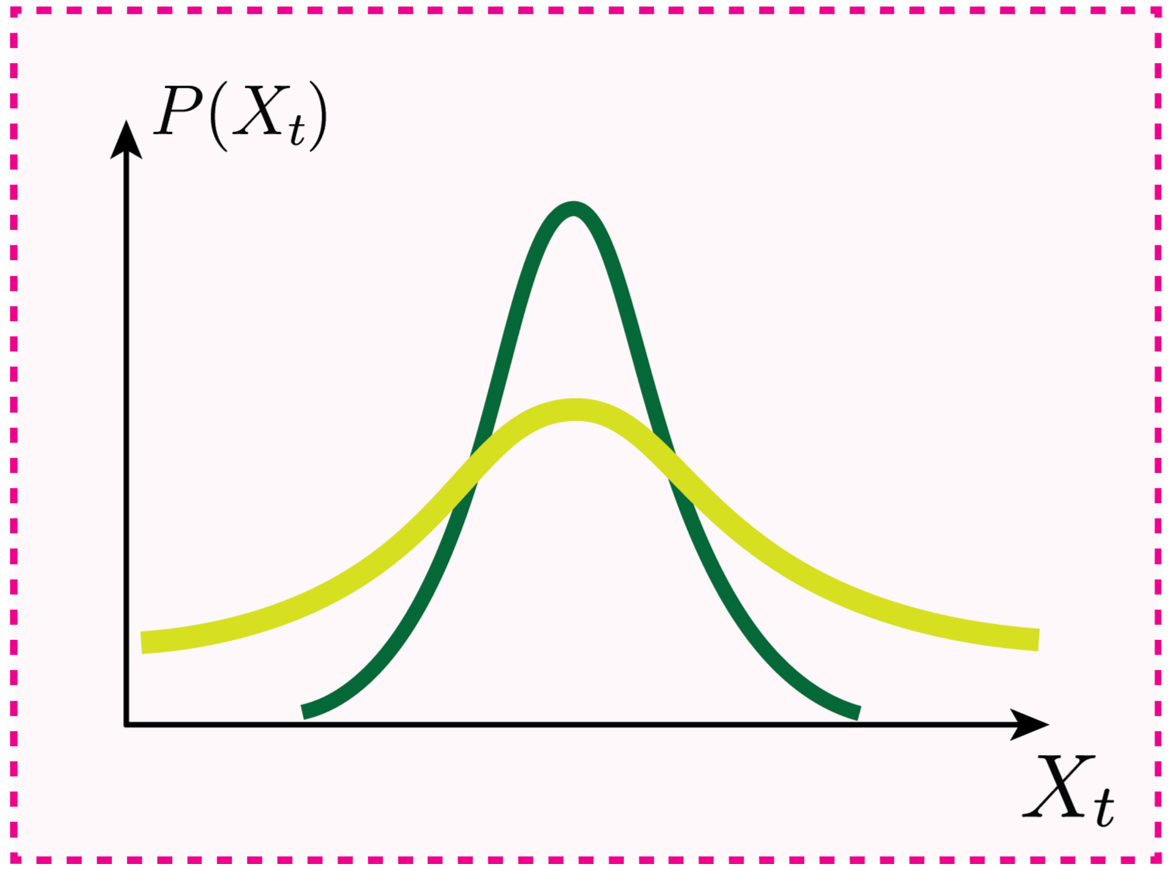


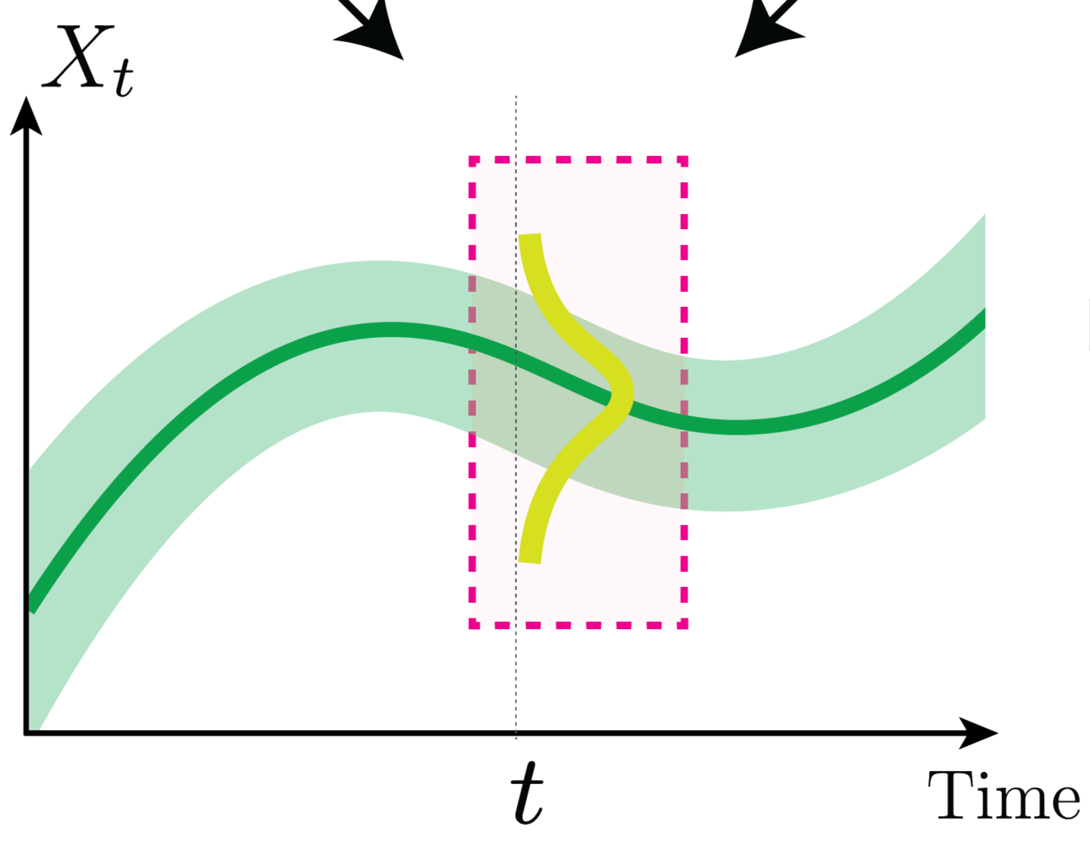
Particle Marginal Metropolis Hastings (PMMH)



Posterior distribution of model parameters

$P(\theta|y_{1:T})$

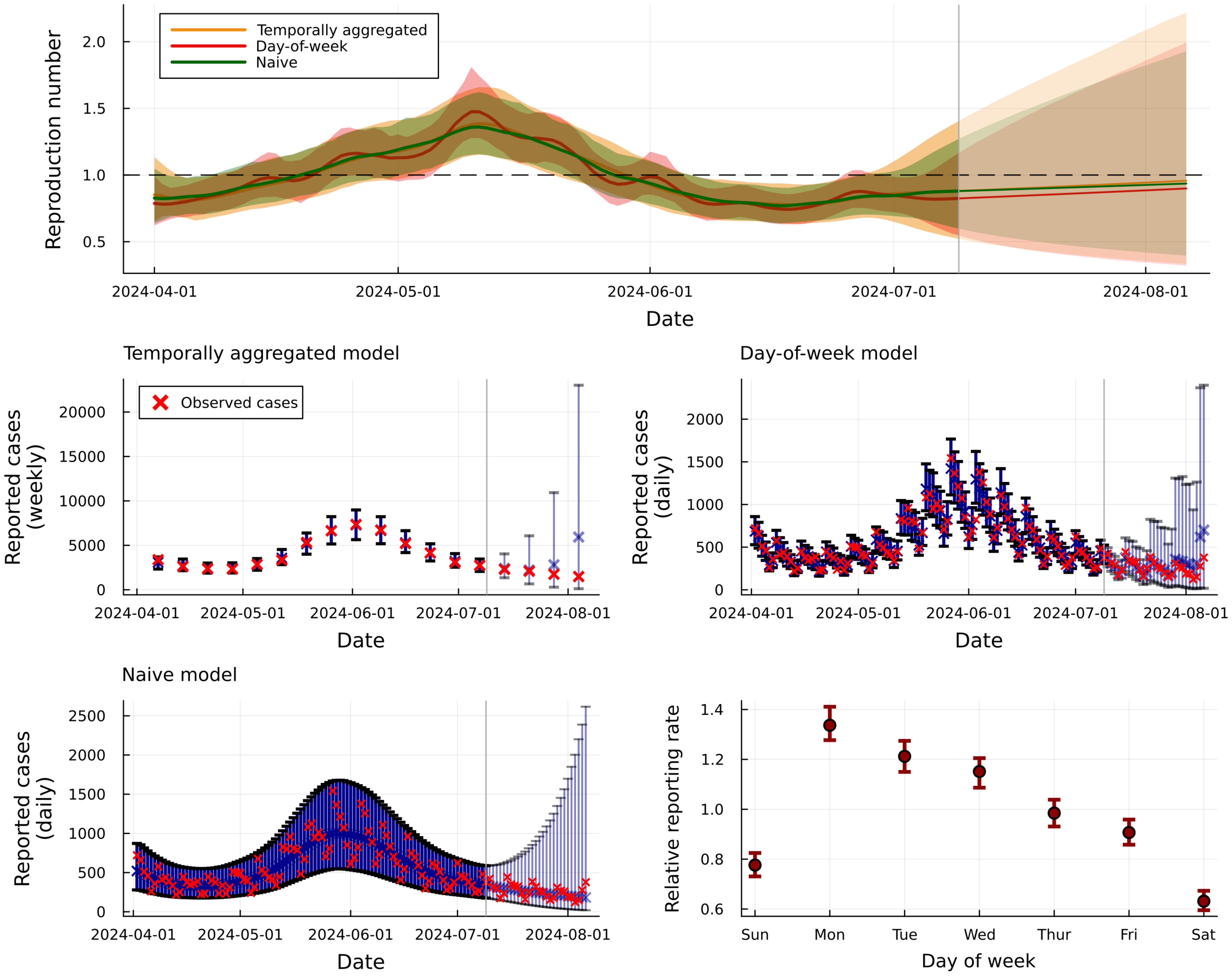


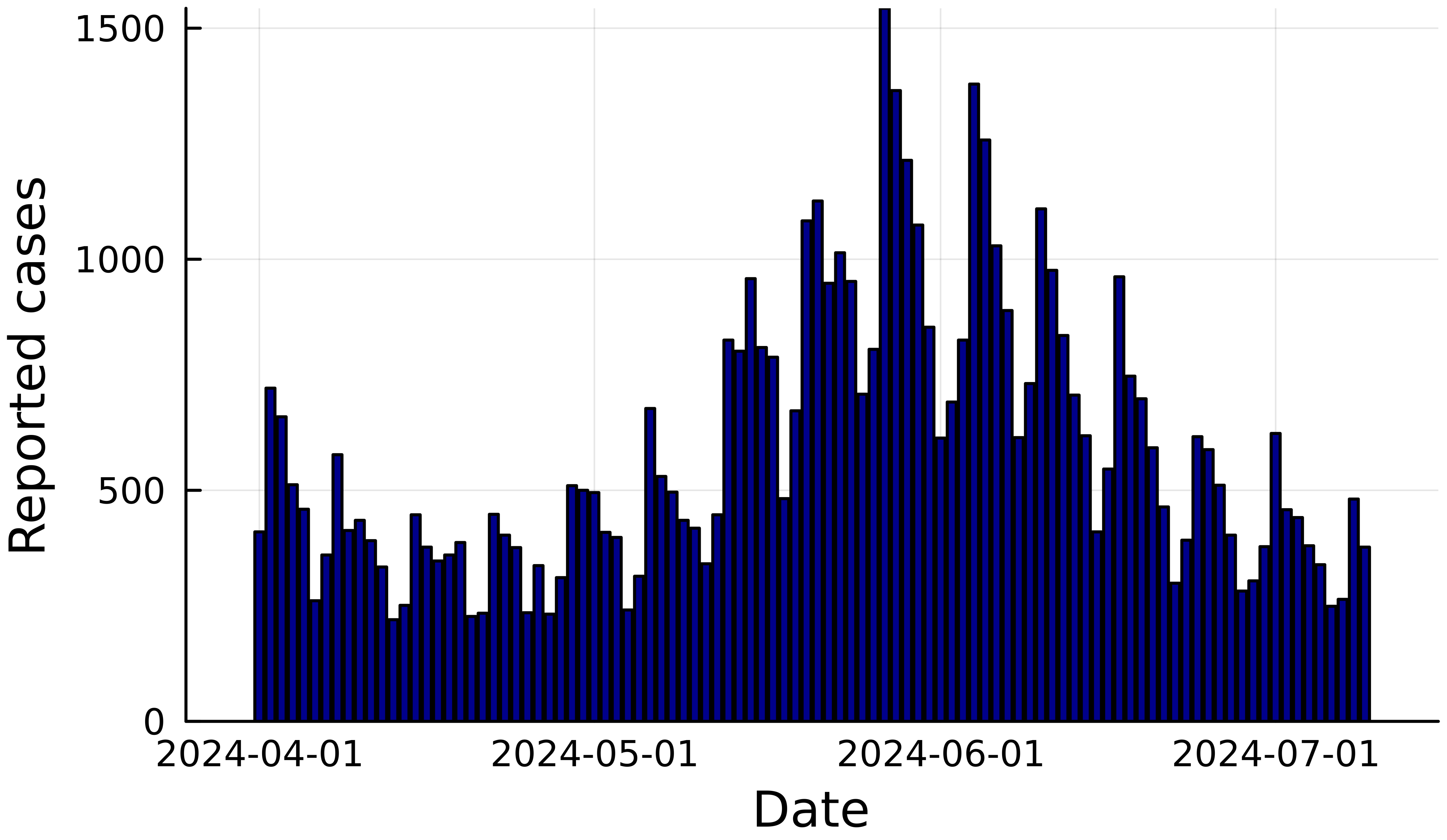


Posterior distribution of hidden states

$P(X_t|y_{1:T})$

Example: Temporal effects





Three observation models:

- Naive
- Day-of-week effect
- Weekly aggregated

Some final thoughts…

Advantages of these methods and models

- ✓ Simple, intuitive, highly flexible
- ✓ Requires no external software
- ✓ No complicated mathematical approximations
- ✓ Produces well-calibrated estimates and predictions

Disadvantages

- x Far less established than existing methods in epi
- x Can struggle with high-dimensional θ
- x Sequential nature is still somewhat restrictive

Website

Preprint

nicholas.steyn@univ.ox.ac.uk