Branching processes forward and backward in time: epidemic prevalence and reconstructed transmission tree.

M. Dávila Felipe, joint work with A. Lambert

Mathematical Models for Epidemiology and Phylogenetics, Lille, May 30 2016

Epidemiology and Phylogenetics

Phylodynamics

"...how pathogen genetic variation, modulated by host immunity, transmission bottlenecks, and epidemic dynamics, determines the wide variety of pathogen phylogenies observed...from individual host to population. We call the melding of immunodynamics, epidemiology, and evolutionary biology required to achieve this synthesis pathogen *phylodynamics*."[Grenfell et al., 2004]

Viral phylodynamics

"...the study of how epidemiological, immunological, and evolutionary processes act and potentially interact to shape viral phylogenies"[Volz et al., 2013]

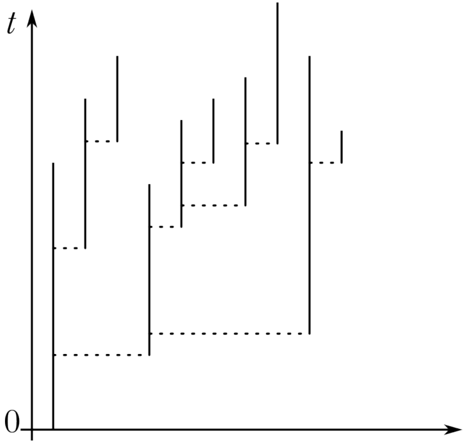
Phylodynamic models

"... population models relating the complex demographics of pathogens to the structure of their phylogenetic trees."[Scarpino, 2016]

Motivation from epidemiology: phylodynamics

- Need to understand: interaction between epidemiological and evolutionary mechanisms
- Few models exist linking pathogen sequence data and prevalence time series
- Many of the existing phylodynamic applications are rooted in a classical Kingman's coalescent framework: poor realism in epidemiology (dense sampling, population size varies stochastically)
- Recent works on modeling and inferring population dynamics from phylogenetic trees (reconstructed from viral sequences), based on birth-death or SIR processes: [Volz et al., 2009, Rasmussen et al., 2011, Stadler et al., 2012, Lambert et al., 2014]

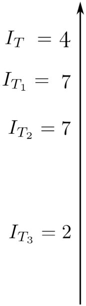
Some conventions



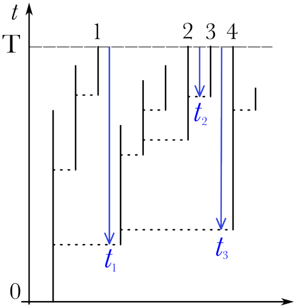
Motivation from the statistical point of view

Data:

Prevalence time series: host population size at deterministic times (not necessarily equidistant)



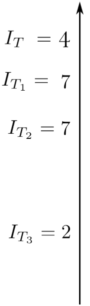
Reconstructed transmission tree: Coalescence times $(t_1, t_2, \dots, t_n)$ estimated from pathogen sequences



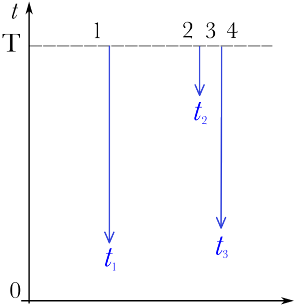
Motivation from the statistical point of view

Data:

Prevalence time series: host population size at deterministic times (not necessarily equidistant)



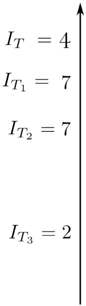
Reconstructed transmission tree: Coalescence times $(t_1, t_2, \ldots, t_n)$ estimated from pathogen sequences



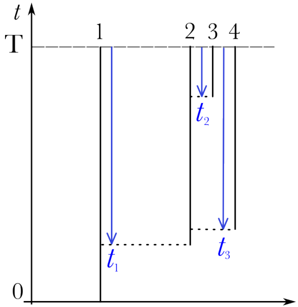
Motivation from the statistical point of view

Data:

Prevalence time series: host population size at deterministic times (not necessarily equidistant)



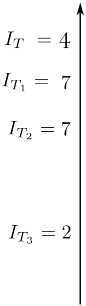
Reconstructed transmission tree: Coalescence times $(t_1, t_2, \dots, t_n)$ estimated from pathogen sequences



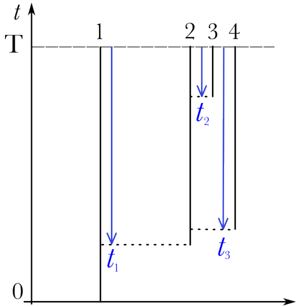
Motivation from the statistical point of view

Data:

Prevalence time series: host population size at deterministic times (not necessarily equidistant)



Reconstructed transmission tree: Coalescence times $(t_1, t_2, \ldots, t_n)$ estimated from pathogen sequences



Goal:

To characterize the joint distribution of $I := (I_t, 0 \leq t \leq T)$ and the reconstructed tree $(t_1, t_2, \ldots, t_n)$

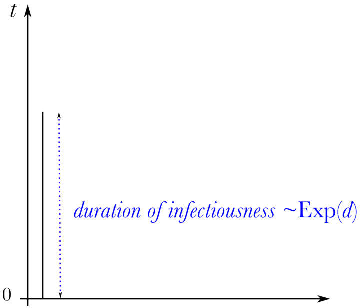
Outline

1. Birth-death models
2. Time reversal for birth-death processes
3. Application to epidemiology
4. Generalization for non-exponential infectiousness periods
5. Ingredients of the proof
6. Conclusions, perspectives and some insight

Outline

1. Birth-death models
2. Time reversal for birth-death processes
3. Application to epidemiology
4. Generalization for non-exponential infectiousness periods
5. Ingredients of the proof
6. Conclusions, perspectives and some insight

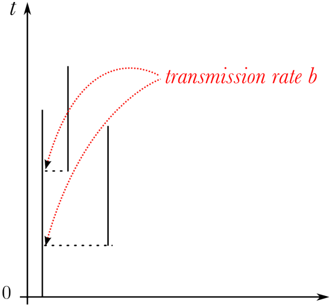
Birth-death (BD) process



Individuals

- have i.i.d. duration of infectiousness $\sim \mathrm{Exp}(d)$
- transmit at constant rate b while infectious
- behave independently from one another

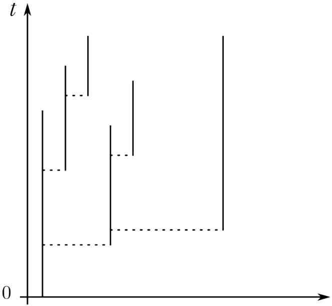
Birth-death (BD) process



Individuals

- have i.i.d. duration of infectiousness $\sim \mathrm{Exp}(d)$
- transmit at constant rate b while infectious
- behave independently from one another

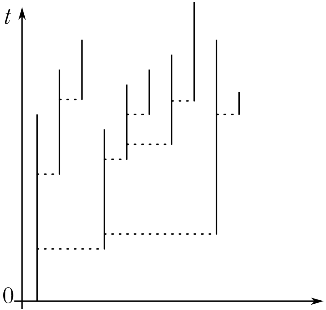
Birth-death (BD) process



Individuals

- have i.i.d. duration of infectiousness $\sim \mathrm{Exp}(d)$
- transmit at constant rate b while infectious
- behave independently from one another

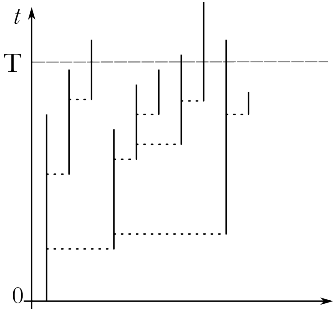
Birth-death (BD) process



Individuals

- have i.i.d. duration of infectiousness $\sim \mathrm{Exp}(d)$
- transmit at constant rate b while infectious
- behave independently from one another

Birth-death (BD) process



Individuals

- have i.i.d. duration of infectiousness $\sim \mathrm{Exp}(d)$
- transmit at constant rate b while infectious
- behave independently from one another

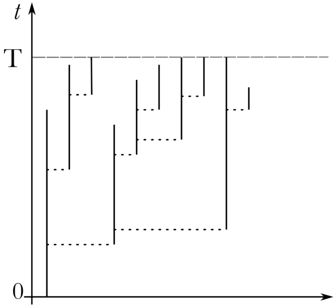
We consider for a fixed time T:

$\mathcal{T}$: the BD tree starting from one ancestor

$\mathcal{T}^{(T)}$: the BD tree truncated up to time T

$(\xi_t(\mathcal{T}), t \geq 0)$: the population size process

Birth-death (BD) process



Individuals

- have i.i.d. duration of infectiousness $\sim \mathrm{Exp}(d)$
- transmit at constant rate b while infectious
- behave independently from one another

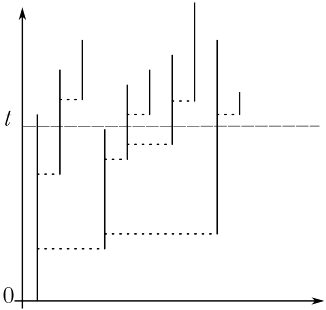
We consider for a fixed time T:

$\mathcal{T}$: the BD tree starting from one ancestor

$\mathcal{T}^{(T)}$: the BD tree truncated up to time T

$(\xi_t(\mathcal{T}), t \geq 0)$: the population size process

Birth-death (BD) process



Individuals

- have i.i.d. duration of infectiousness $\sim \mathrm{Exp}(d)$
- transmit at constant rate b while infectious
- behave independently from one another

We consider for a fixed time T:

$\mathcal{T}$: the BD tree starting from one ancestor

$\mathcal{T}^{(T)}$: the BD tree truncated up to time T

$(\xi_t(\mathcal{T}), t \geq 0)$: the population size process

Random forests

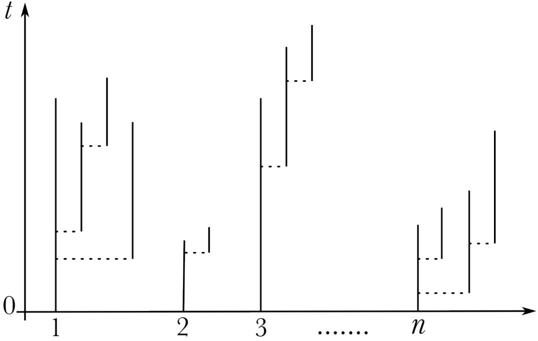
Forest $\mathcal{F}$:

A finite sequence of i.i.d BD trees $(\mathcal{T}_1, \ldots, \mathcal{T}_n)$

Random forests

Forest $\mathcal{F}$:

A finite sequence of i.i.d BD trees $(\mathcal{T}_1, \ldots, \mathcal{T}_n)$



Random forests

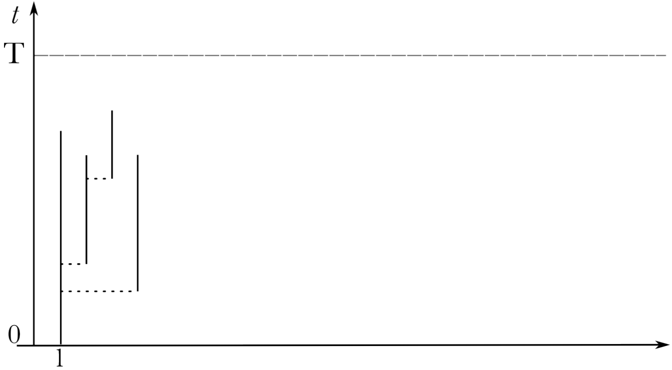
Forest $\mathcal{F}^*$:

A sequence of i.i.d. BD trees stopped at the first tree that survives up until time T

Random forests

Forest $\mathcal{F}^*$:

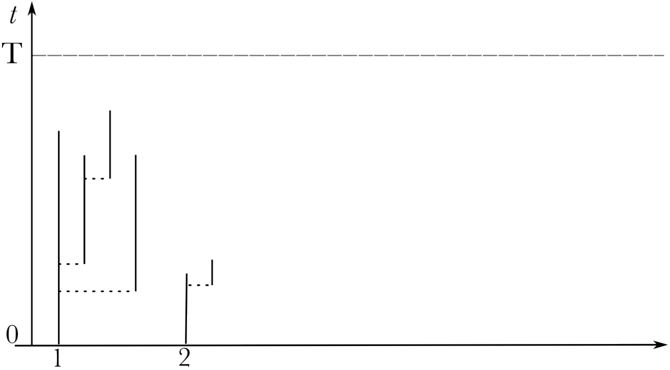
A sequence of i.i.d. BD trees stopped at the first tree that survives up until time T



Random forests

Forest $\mathcal{F}^*$:

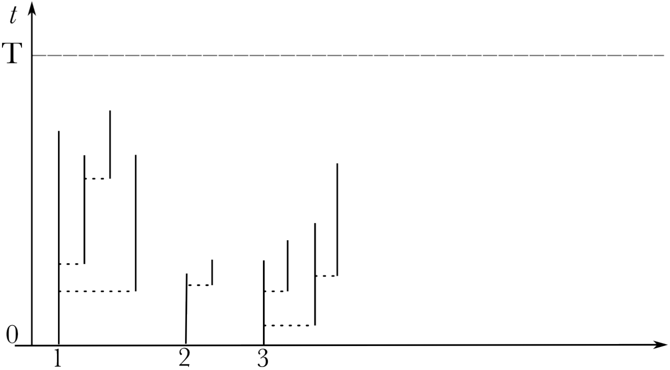
A sequence of i.i.d. BD trees stopped at the first tree that survives up until time T



Random forests

Forest $\mathcal{F}^*$:

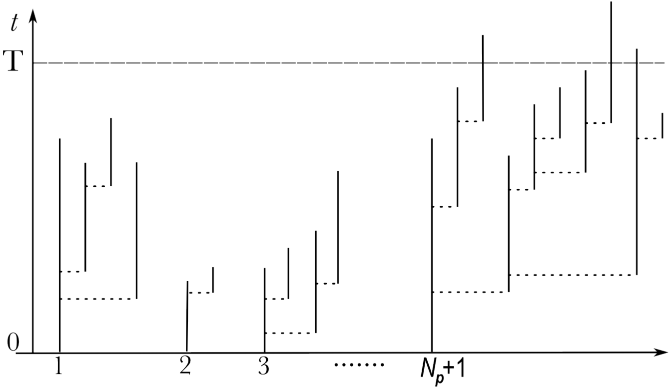
A sequence of i.i.d. BD trees stopped at the first tree that survives up until time T



Random forests

Forest $\mathcal{F}^*$:

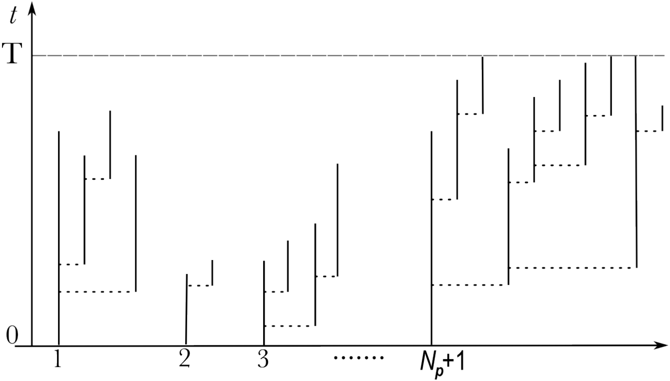
A sequence of i.i.d. BD trees stopped at the first tree that survives up until time T



Random forests

Forest $\mathcal{F}^*$:

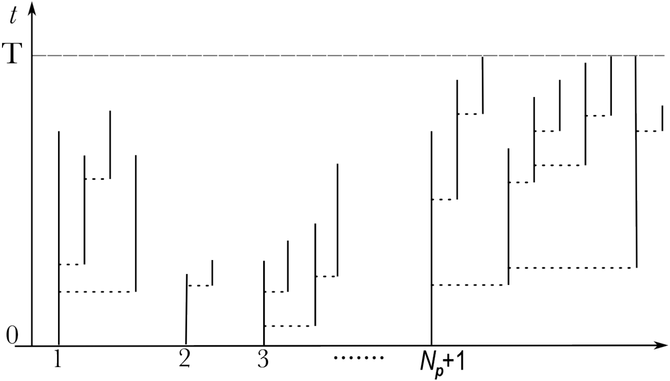
A sequence of i.i.d. BD trees stopped at the first tree that survives up until time T



Random forests

Convention for $\mathcal{F}^*$:

A forest stopped at the first surviving tree (up to time T)



Random forests

Forest $\mathcal{F}$:

A finite sequence of i.i.d BD trees $(\mathcal{T}_1, \ldots, \mathcal{T}_n)$

Forest $\mathcal{F}^*$:

A sequence of i.i.d. BD trees stopped at the first tree that survives up until time T

For any forest $\mathcal{F}$, the population size process is denoted by,

$$\left(\xi_t\left(\mathcal{F}\right), t \geq 0\right)$$

Outline

1. Birth-death models
2. Time reversal for birth-death processes
3. Application to epidemiology
4. Generalization for non-exponential infectiousness periods
5. Ingredients of the proof
6. Conclusions, perspectives and some insight

Time-reversal duality

Fix $b \geq d$ and $T > 0$

$\mathcal{F}^*$:= Forest stopped at 1st surviving tree
with birth and death parameters (b, d)

Time-reversal duality

Fix $b \geq d$ and $T > 0$

$\mathcal{F}^*$:= Forest stopped at 1st surviving tree with birth and death parameters (b, d)

$\widetilde{\mathcal{F}}^*$:= Forest stopped at 1st surviving tree with birth and death parameters (d, b)

Time-reversal duality

Fix $b \geq d$ and $T > 0$

$\mathcal{F}^*$:= Forest stopped at 1st surviving tree with birth and death parameters (b, d)

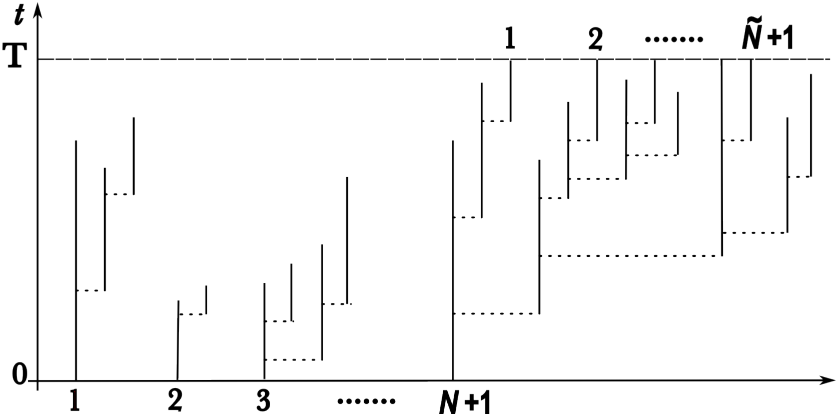
$\widetilde{\mathcal{F}}^*$:= Forest stopped at 1st surviving tree with birth and death parameters (d, b)

Theorem [D.F. and Lambert, 2015]

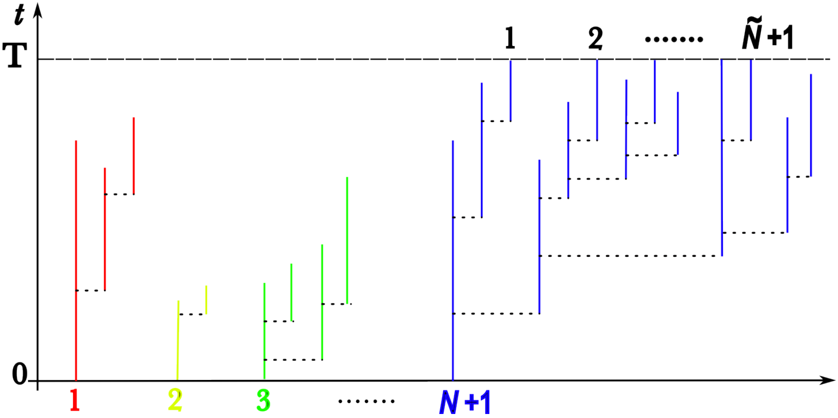
We have the following identity in distribution,

$$\left(\xi_{T-t}\left(\mathcal{F}^*\right), 0 \leq t \leq T\right) \overset{d}{=} \left(\xi_t\left(\widetilde{\mathcal{F}}^*\right), 0 \leq t \leq T\right)$$

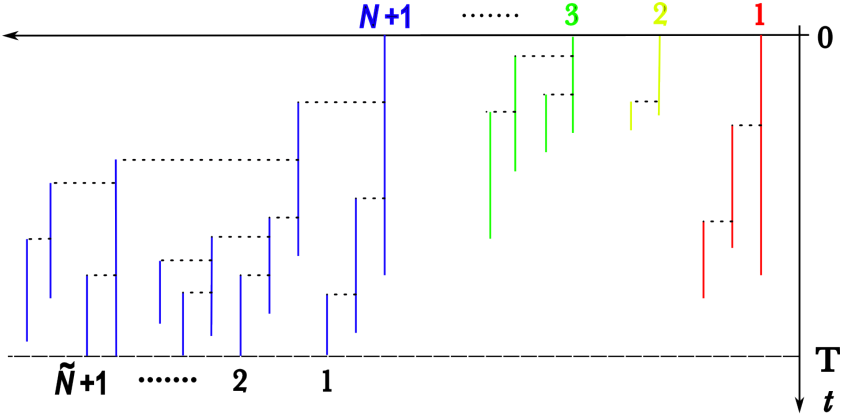
Space-time-reversal duality



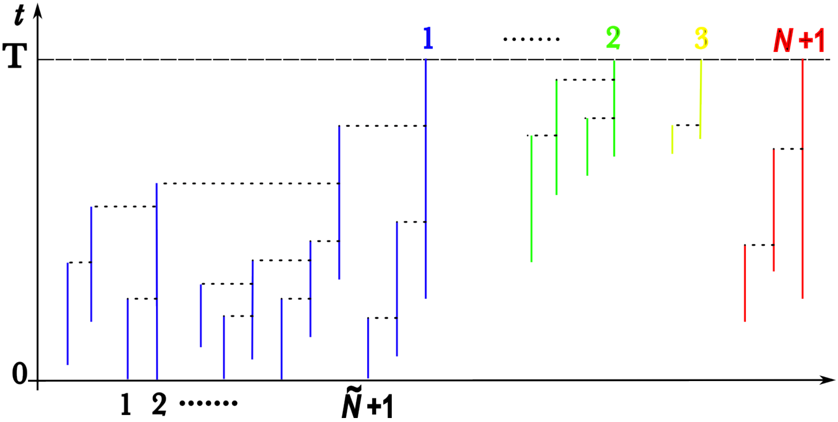
Space-time-reversal duality



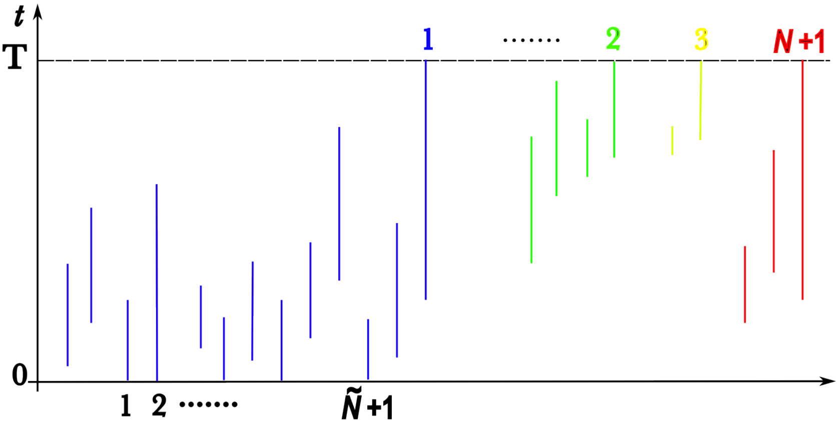
Space-time-reversal duality



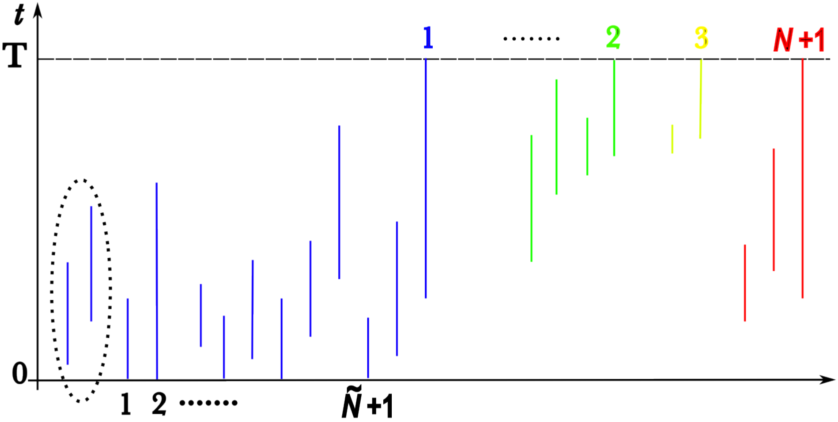
Space-time-reversal duality



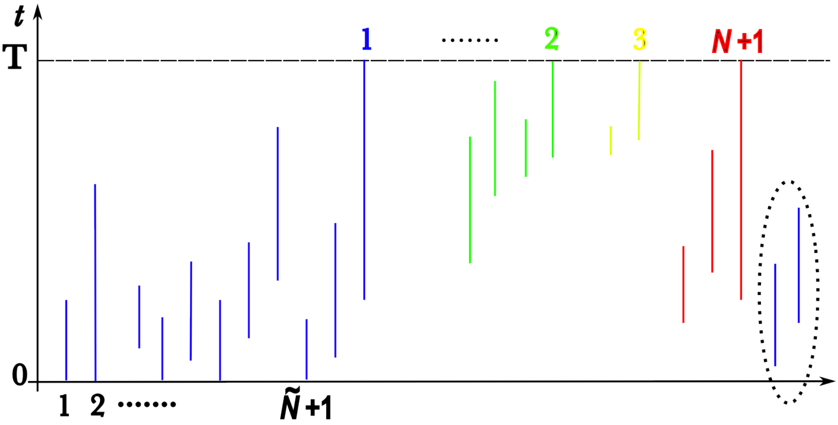
Space-time-reversal duality



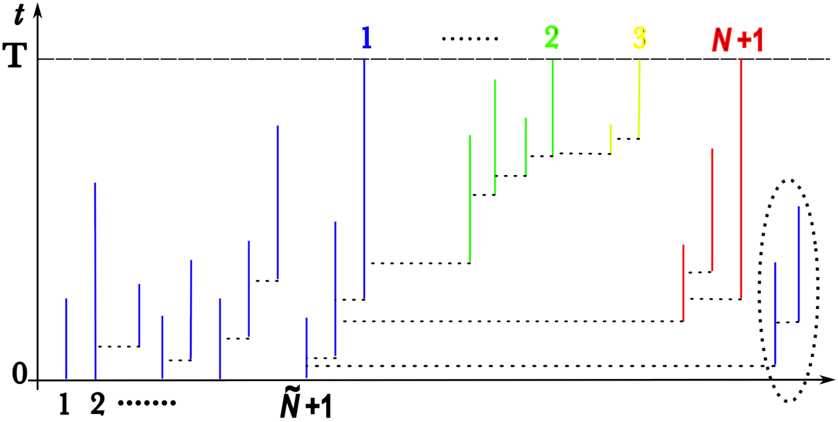
Space-time-reversal duality



Space-time-reversal duality



Space-time-reversal duality



Outline

1. Birth-death models
2. Time reversal for birth-death processes
3. **Application to epidemiology**
4. Generalization for non-exponential infectiousness periods
5. Ingredients of the proof
6. Conclusions, perspectives and some insight

Phylodynamics

Goal:

To infer the parameters on the model from:

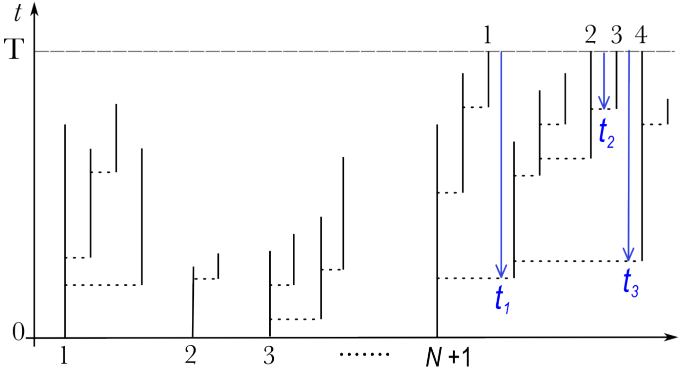
- prevalence time series: $(I_{T_\mathbf{0}}, I_{T_\mathbf{1}} \dots, I_{T_N})$, i.e. the host **population size** at deterministic times
  $T_0 > T_1 \dots > T_N > 0$, where $T_0 = T$ is present time
- sequence data $\Longrightarrow$ the **reconstructed tree**

Conditional on the reduced tree: applications to epidemiology

We want to characterize the population size process conditional on the coalescence times between extant hosts at present time T to be $t_1, t_2, \ldots, t_n$

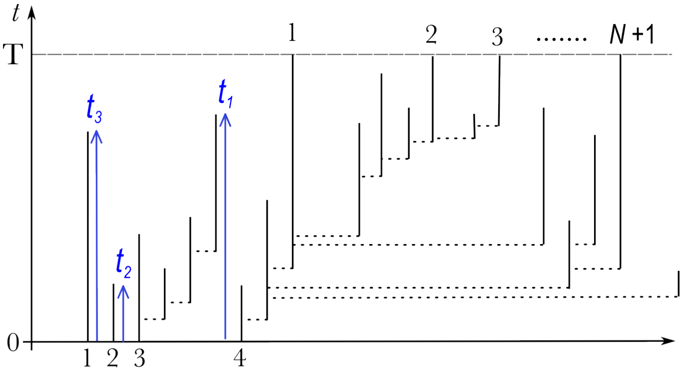
Conditional on the reduced tree: applications to epidemiology

We want to characterize the population size process conditional on the coalescence times between extant hosts at present time T to be $t_1, t_2, \dots, t_n$

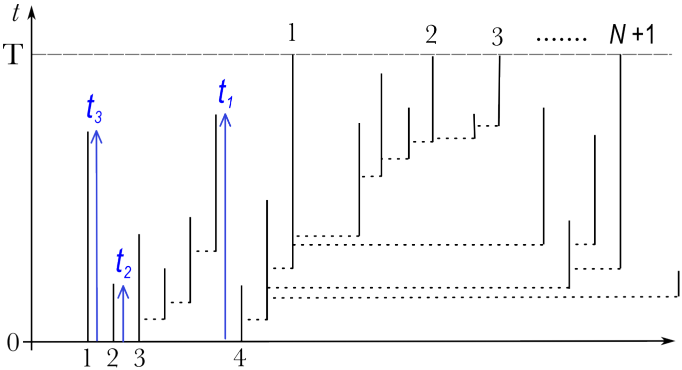


Conditional on the reduced tree: applications to epidemiology

When we return the time, thanks to the duality property, coalescence times become life durations



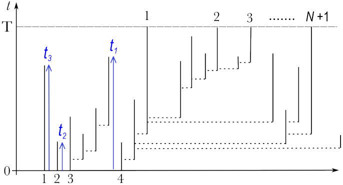
Conditional on the reduced tree: applications to epidemiology



Theorem [D.F. and Lambert, 2015]

The population size process conditional on the coalescence times to be $t_1, \dots, t_n$, backward in time, is that of a sum of n BD trees, each conditioned on dying out at t_i for $1 \leq i \leq n$, plus an additional tree conditioned on surviving up until time T.

Conditional on the reduced tree: applications to epidemiology



Theorem [D.F. and Lambert, 2015]

Let $b \geq d$, $\mathbb{P} := \mathbb{P}_{b,d}$, $\widetilde{\mathbb{P}} := \mathbb{P}_{d,b}$ and $\mathcal{F}^*$ be a forest stopped at 1st surviving tree with parameters b and d. Let $(H_i)_{1\leq i\leq N}$ be the coalescence times from individuals at T and define

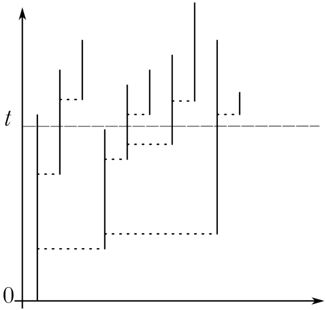
$$\widetilde{\mathbb{P}}^{(i)} := \widetilde{\mathbb{P}}(\cdot\,|\,T_{\mathrm{Ext}} = t_i), \quad \forall 1 \leq i \leq N \qquad \text{and} \qquad \widetilde{\mathbb{P}}^{(N+1)} := \widetilde{\mathbb{P}}(\cdot\,|\,T_{\mathrm{Ext}} > T).$$

Then

$$\left(\xi_{T-t}(\mathcal{F}^*),\ \mathbb{P}(\cdot|H_i = t_i, 1 \leq i \leq N)\right) \;=\; \left(\sum_{i=1}^{N+1} \xi_t\left(\mathcal{T}_i\right),\ (\widetilde{\mathbb{P}}^{(1)} * \ldots * \widetilde{\mathbb{P}}^{(N+1)})(\cdot)\right)$$

Outline

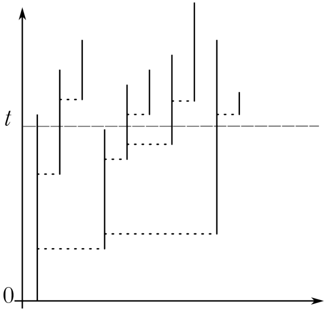
Splitting trees



Individuals

- have i.i.d. infectiousness durations with general distribution
- transmit at constant rate b while infectious
- behave independently from one another

Splitting trees

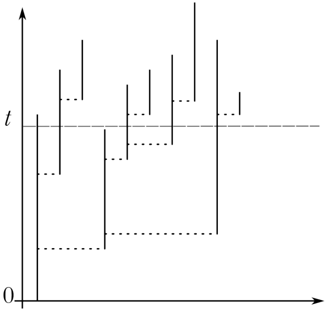


Individuals

- have i.i.d. infectiousness durations with general distribution
- transmit at constant rate *b* while infectious
- behave independently from one another

A splitting tree is characterized by a σ-finite measure Π on $(0, \infty)$ satisfying $\int_{(0,\infty)} (1 \wedge r)\Pi(\mathrm{d}r) < \infty$ (the *lifespan measure*).

Splitting trees



Individuals

- have i.i.d. infectiousness durations with general distribution
- transmit at constant rate *b* while infectious
- behave independently from one another

A splitting tree is characterized by a σ-finite measure Π on $(0, \infty)$ satisfying $\int_{(0,\infty)}(1 \wedge r)\Pi(\mathrm{d}r) < \infty$ (the *lifespan measure*).

We consider Π finite with mass b: individuals give birth (transmit) at rate b and have life (infectiousness) durations distributed as $\Pi(\cdot)/b$.

Time reversal duality for splitting trees

Define for Π:

- The Laplace exponent: $\psi(\lambda) := \lambda - \int_0^\infty \left(1 - \mathrm{e}^{-\lambda r}\right) \Pi(\mathrm{d}r),\ \lambda \geq 0$
- η the Malthusian parameter of the epidemic (largest root of ψ)
- $m := \int_0^\infty r\Pi(\mathrm{d}r)$
- A new measure $\widetilde{\Pi}(\mathrm{d}r) := \mathrm{e}^{-\eta r}\Pi(\mathrm{d}r)$

Time reversal duality for splitting trees

Define for Π:

- The Laplace exponent: $\psi(\lambda) := \lambda - \int_0^\infty \left(1 - \mathrm{e}^{-\lambda r}\right) \Pi(\mathrm{d}r),\ \lambda \geq 0$
- η the Malthusian parameter of the epidemic (largest root of ψ)
- $m := \int_0^\infty r\Pi(\mathrm{d}r)$
- A new measure $\widetilde{\Pi}(\mathrm{d}r) := \mathrm{e}^{-\eta r}\Pi(\mathrm{d}r)$

The scale function W:

The unique continuous function $W : [0, +\infty) \to [0, +\infty)$, characterized by its Laplace transform,

$$\int_0^{+\infty} \mathrm{e}^{-\lambda x} W(x) = \frac{1}{\psi(\lambda)}, \qquad \lambda > \eta$$

Time reversal duality for splitting trees

Define for Π:

- The Laplace exponent: $\psi(\lambda) := \lambda - \int_0^\infty \left(1 - \mathrm{e}^{-\lambda r}\right) \Pi(\mathrm{d}r),\ \lambda \geq 0$
- η the Malthusian parameter of the epidemic (largest root of ψ)
- $m := \int_0^\infty r\Pi(\mathrm{d}r)$
- A new measure $\widetilde{\Pi}(\mathrm{d}r) := \mathrm{e}^{-\eta r}\Pi(\mathrm{d}r)$

The scale function W:

The unique continuous function $W : [0, +\infty) \to [0, +\infty)$, characterized by its Laplace transform,

$$\int_0^{+\infty} \mathrm{e}^{-\lambda x} W(x) = \frac{1}{\psi(\lambda)}, \qquad \lambda > \eta$$

Define:

$$\gamma = \frac{1}{W(T)} \qquad \widetilde{\gamma} = \frac{1}{\widetilde{W}(T)}$$

Time reversal duality for splitting trees

Forest $\mathcal{F}^p$:

A sequence of independent splitting trees $\left(\mathcal{T}_1, \ldots, \mathcal{T}_{N_p}, \mathcal{T}_{N_p+1}^{(T)}\right) \perp\!\!\!\perp N_p$, where,

- $\mathcal{T}_1, \ldots \mathcal{T}_{N_p}$: are i.i.d. conditioned on extinction before T
- $\mathcal{T}_{N_p+1}$: is conditionned on survival up until time T
- N_p: is a geometric random variable with $\mathbb{P}(N_p = k) = (1-p)^k p$, $k \geq 0$

Time reversal duality for splitting trees

Forest $\mathcal{F}^p$:

A sequence of independent splitting trees $\left(\mathcal{T}_1, \ldots, \mathcal{T}_{N_p}, \mathcal{T}_{N_p+1}^{(T)}\right) \perp\!\!\!\perp N_p$, where,

- $\mathcal{T}_1, \ldots \mathcal{T}_{N_p}$: are i.i.d. conditioned on extinction before T
- $\mathcal{T}_{N_p+1}$: is conditionned on survival up until time T
- N_p: is a geometric random variable with $\mathbb{P}(N_p = k) = (1-p)^k p$, $k \geq 0$

$\mathcal{F}_\top^p, \mathcal{F}_\perp^p$:

$\sim \mathcal{F}^p$, but lifetimes of the ancestors have a specific distribution $\mu_\top, \mu_\perp$, $\neq$ from $\Pi(\cdot)/b$

Time reversal duality for splitting trees

Forest $\mathcal{F}^p$:

A sequence of independent splitting trees $\left(\mathcal{T}_1, \dots, \mathcal{T}_{N_p}, \mathcal{T}^{(T)}_{N_p+1}\right) \perp\!\!\!\perp N_p$, where,

- $\mathcal{T}_1, \dots \mathcal{T}_{N_p}$: are i.i.d. conditioned on extinction before T
- $\mathcal{T}_{N_p+1}$: is conditionned on survival up until time T
- N_p: is a geometric random variable with $\mathbb{P}(N_p = k) = (1-p)^k p$, $k \geq 0$

$\mathcal{F}^p_\top, \mathcal{F}^p_\perp$:

$\sim \mathcal{F}^p$, but lifetimes of the ancestors have a specific distribution $\mu_\top, \mu_\perp$, $\neq$ from $\Pi(\cdot)/b$

Lemma (Supercritical case: $m \geq 1$, $\eta > 0$) [D.F. and Lambert, 2015]

$\mathcal{F}^{\widetilde{\gamma}}_\perp$ = a forest stopped at first surviving tree with ancestors and individuals $\sim (\perp, \Pi)$

$\widetilde{\mathcal{F}}^{\gamma}_\top$ = a forest stopped at first surviving tree with ancestors and individuals $(\top, \widetilde{\Pi})$.

Time reversal duality for splitting trees

Lemma (Supercritical case: $m \geq 1$, $\eta > 0$) [D.F. and Lambert, 2015]

$\mathcal{F}_{\perp}^{\widetilde{\gamma}}$ = a forest stopped at first surviving tree with ancestors and individuals $\sim (\perp, \Pi)$

$\widetilde{\mathcal{F}}_{\top}^{\gamma}$ = a forest stopped at first surviving tree with ancestors and individuals $(\top, \widetilde{\Pi})$.

Theorem [D.F. and Lambert, 2015]

We have the following identity in distribution,

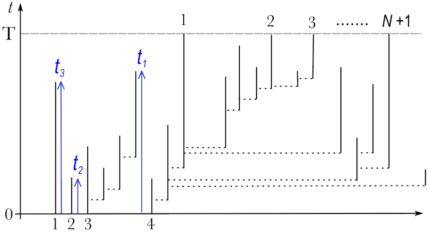
$$\left(\xi_{T-t}\left(\mathcal{F}_{\perp}^{\widetilde{\gamma}}\right), 0 \leq t \leq T\right) \overset{d}{=} \left(\xi_{t}\left(\widetilde{\mathcal{F}}_{\top}^{\gamma}\right), 0 \leq t \leq T\right)$$

In the subcritical and critical cases (i.e. $m \leq 1$), then,

$$\left(\xi_{T-t}\left(\mathcal{F}_{\perp}^{\gamma}\right), 0 \leq t \leq T\right) \overset{d}{=} \left(\xi_{t}\left(\mathcal{F}_{\top}^{\gamma}\right), 0 \leq t \leq T\right)$$

and actually in this case $\mu_{\perp}(\mathrm{d}r) = \mu_{\top}(\mathrm{d}r) = \dfrac{\overline{\Pi}(r)}{m}\mathrm{d}r$.

Conditional on the reduced tree: applications to epidemiology



Theorem [D.F. and Lambert, 2015]

Let $\mathcal{F}_\perp^{\tilde\gamma}$ and $(H_i)_{i\geq 1}$ the coalescence times from individuals at T. Define

$$\widetilde{\mathbb{P}}_{\top,i} := \widetilde{\mathbb{P}}_{\top}(\cdot|\, T_{\mathrm{Ext}} = t_i),\ \ \forall i \geq 1 \qquad \text{and} \qquad \widetilde{\mathbb{P}}_{\top,*} := \widetilde{\mathbb{P}}_{\top}(\cdot|\, T_{\mathrm{Ext}} > T).$$

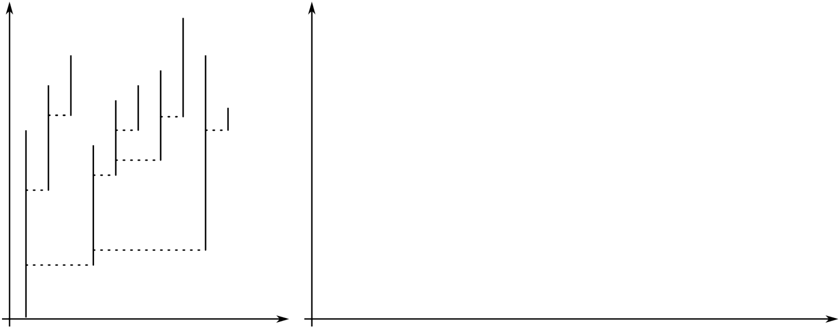
Then

$$\left(\xi_{T-t}(\mathcal{F}_\perp^{\tilde\gamma}),\ \mathbb{P}(\cdot|H_i = t_i, 1 \leq i \leq N)\right) \ = \ \left(\sum_{i=1}^{N+1} \xi_t\left(\mathcal{T}_i\right),\ (\widetilde{\mathbb{P}}_{\top,1} * \ldots * \widetilde{\mathbb{P}}_{\top,*})(\cdot)\right)$$

Outline

1. Birth-death models
2. Time reversal for birth-death processes
3. Application to epidemiology
4. Generalization for non-exponential infectiousness periods
5. Ingredients of the proof
6. Conclusions, perspectives and some insight

The jumping chronological contour process [Lambert, 2010]

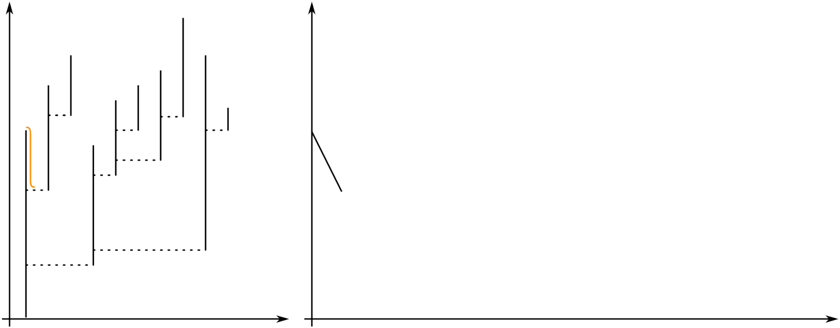


Example of a finite splitting tree and its contour process[1]

---

[1] Figures from C. Delaporte - Aussois 2013

The jumping chronological contour process [Lambert, 2010]

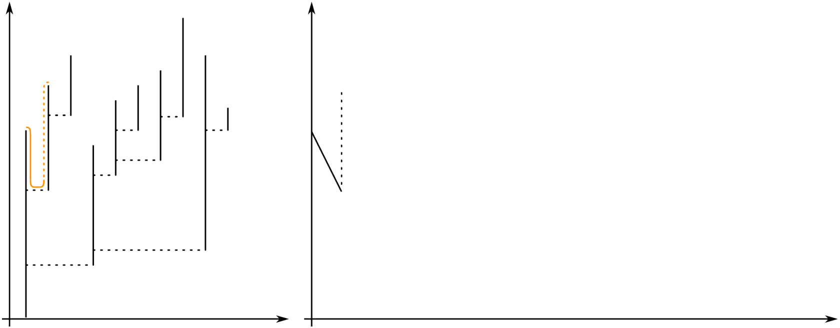


Example of a finite splitting tree and its contour process[1]

---

[1] Figures from C. Delaporte - Aussois 2013

The jumping chronological contour process [Lambert, 2010]

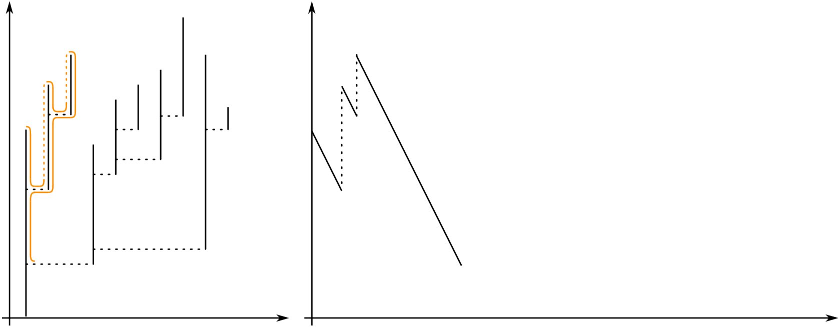


Example of a finite splitting tree and its contour process[1]

---

[1] Figures from C. Delaporte - Aussois 2013

The jumping chronological contour process [Lambert, 2010]

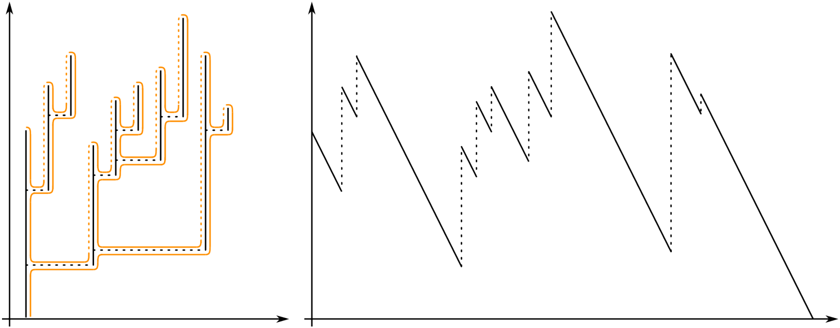


Example of a finite splitting tree and its contour process[1]

---

[1] Figures from C. Delaporte - Aussois 2013

The jumping chronological contour process [Lambert, 2010]



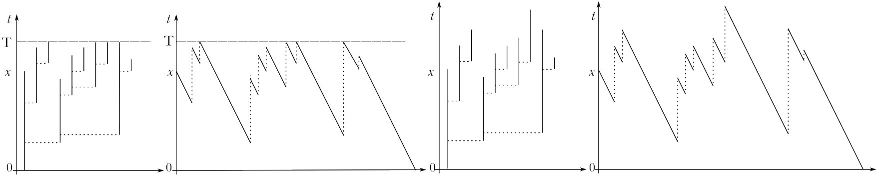
Example of a finite splitting tree and its contour process[1]

---

[1] Figures from C. Delaporte - Aussois 2013

The contour of splitting trees is a Lévy process

Let Y be a a finite variation Lévy process with Lévy measure Π and drift -1.



Theorem [Lambert, 2010]

Conditional on the lifespan of the ancestor to be x, the contour of $\mathcal{T}^{(T)}$, is distributed as Y, started at $x \wedge T$, reflected below T and killed upon hitting 0.

The contour of $\mathcal{T}$, conditional on extinction, has the law of Y started at x, conditioned on, and killed upon hitting 0.

Time reversal duality for spectrally positive Lévy processes

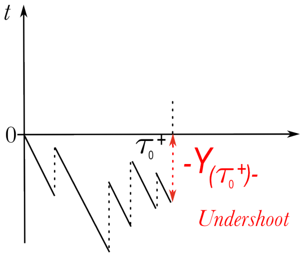
Theorem [Bertoin, 1992]

The excursion measure has the following property of invariance under time reversal: under $P_0\left(\cdot\middle|-Y_{(\tau_{\mathbf{0}}^+)-}=u\right)$ the reverted excursion, $\left(-Y_{(\tau_{\mathbf{0}}-t)-},0\leq t<\tau_0\right)$ has the same distribution that $(Y_t,0\leq t<\tau_0)$ under $P_u\left(\cdot|\tau_0<+\infty\right)$.

Time reversal duality for spectrally positive Lévy processes

Theorem [Bertoin, 1992]

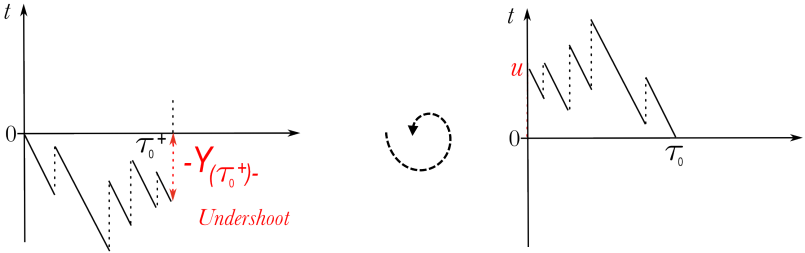
The excursion measure has the following property of invariance under time reversal: under $P_0\left(\cdot\middle|-Y_{(\tau_{\mathbf{0}}^+)-}=u\right)$ the reverted excursion, $\left(-Y_{(\tau_{\mathbf{0}}-t)-}, 0\leq t<\tau_0\right)$ has the same distribution that $(Y_t, 0\leq t<\tau_0)$ under $P_u\left(\cdot|\tau_0<+\infty\right)$.



Time reversal duality for spectrally positive Lévy processes

Theorem [Bertoin, 1992]

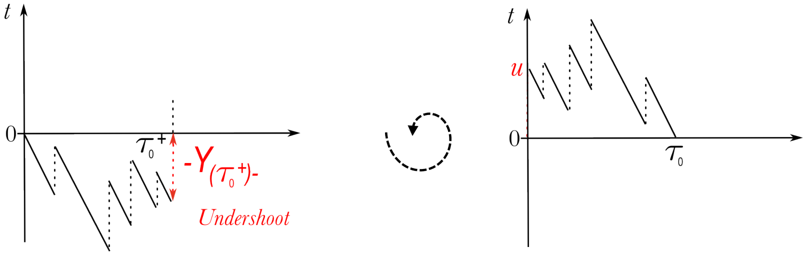
The excursion measure has the following property of invariance under time reversal: under $P_0\left(\cdot\,\middle|\,-Y_{(\tau_0^+)-}=u\right)$ the reverted excursion, $\left(-Y_{(\tau_0-t)-}, 0\leq t<\tau_0\right)$ has the same distribution that $(Y_t, 0\leq t<\tau_0)$ under $P_u\left(\cdot|\tau_0<+\infty\right)$.



Time reversal duality for spectrally positive Lévy processes

Theorem [Bertoin, 1992]

The excursion measure has the following property of invariance under time reversal: under $P_0\left(\cdot\middle|-Y_{(\tau_{\mathbf{0}}^{+})-}=u\right)$ the reverted excursion, $\left(-Y_{(\tau_{\mathbf{0}}-t)-},0\leq t<\tau_0\right)$ has the same distribution that $(Y_t,0\leq t<\tau_0)$ under $P_u\left(\cdot|\tau_0<+\infty\right)$.



From [Bertoin, 1996, Lambert, 2010]:

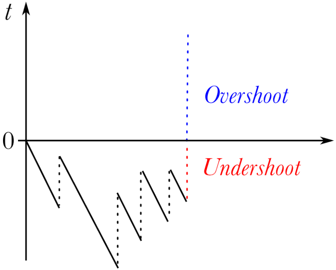
In the **supercritical** case, under $P\left(\cdot|\tau_0<+\infty\right)$, Y is **subcritical** with Lévy measure $\widetilde{\Pi}$.

Undershoot and overshoot

$\mathcal{F}^p_\top, \mathcal{F}^p_\perp$:

Lifetimes of the ancestors have a specific distribution, different from $\Pi(\cdot)/b$:

The undershoot and overshoot at 0 of an excursion starting at 0 and conditional on $\tau_0^+ < +\infty$, are distributed as follows,



Overshoot ($\perp$): $\sim \dfrac{\mathrm{e}^{\eta v}\overline{\widetilde{\Pi}}(v)\mathrm{d}v}{m \wedge 1}$

Undershoot ($\top$): $\sim \dfrac{\mathrm{e}^{-\eta u}\overline{\Pi}(u)\mathrm{d}u}{m \wedge 1}$

Time reversal duality for splitting trees

Define for a Lévy measure Π:

- The Laplace exponent: $\psi(\lambda) := \lambda - \int_0^\infty \left(1 - \mathrm{e}^{-\lambda r}\right) \Pi(\mathrm{d}r),\ \lambda \geq 0$
- η the largest root of ψ
- A new measure $\widetilde{\Pi}(\mathrm{d}r) := \mathrm{e}^{-\eta r}\Pi(\mathrm{d}r)$
- $\tau_A = \in \{t \geq 0 : Y_t \in A\}$ the first hitting time of the real Borel set A

Time reversal duality for splitting trees

Define for a Lévy measure Π:

- The Laplace exponent: $\psi(\lambda) := \lambda - \int_0^\infty \left(1 - \mathrm{e}^{-\lambda r}\right) \Pi(\mathrm{d}r),\ \lambda \geq 0$
- η the largest root of ψ
- A new measure $\widetilde{\Pi}(\mathrm{d}r) := \mathrm{e}^{-\eta r}\Pi(\mathrm{d}r)$
- $\tau_A =\in \{t \geq 0 : Y_t \in A\}$ the first hitting time of the real Borel set A

The scale function W:

The unique continuous function $W : [0, +\infty) \to [0, +\infty)$, characterized by its Laplace transform,

$$\int_0^{+\infty} \mathrm{e}^{-\lambda x} W(x) = \frac{1}{\psi(\lambda)}, \qquad \lambda > \eta$$

Time reversal duality for splitting trees

Define for a Lévy measure Π:

- The Laplace exponent: $\psi(\lambda) := \lambda - \int_0^\infty \left(1 - \mathrm{e}^{-\lambda r}\right) \Pi(\mathrm{d}r),\ \lambda \geq 0$
- η the largest root of ψ
- A new measure $\widetilde{\Pi}(\mathrm{d}r) := \mathrm{e}^{-\eta r}\Pi(\mathrm{d}r)$
- $\tau_A =\in \{t \geq 0 : Y_t \in A\}$ the first hitting time of the real Borel set A

The scale function W:

The unique continuous function $W : [0, +\infty) \to [0, +\infty)$, characterized by its Laplace transform,

$$\int_0^{+\infty} \mathrm{e}^{-\lambda x} W(x) = \frac{1}{\psi(\lambda)}, \qquad \lambda > \eta$$

Why?:

$$\gamma = \frac{1}{W(T)} \qquad \widetilde{\gamma} = \frac{1}{\widetilde{W}(T)}$$

Excursions of spectrally positive Lévy processes

Two-sided-exit problem [Bertoin, 1996]

For $0 \leq x \leq a$,

$$P_x\left(\tau_0 < \tau_a^+\right) = \frac{W(a-x)}{W(a)}$$

Excursions of spectrally positive Lévy processes

Two-sided-exit problem [Bertoin, 1996]

For $0 \leq x \leq a$,

$$P_x\left(\tau_0 < \tau_a^+\right) = \frac{W(a-x)}{W(a)}$$

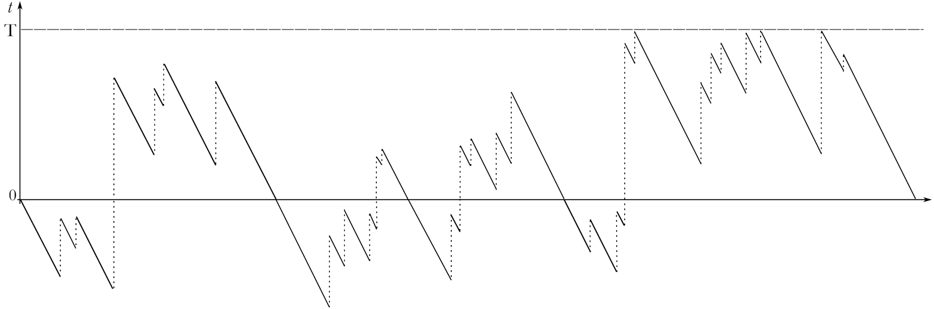
Supercritica case:

$$\begin{aligned} \gamma &= \frac{1}{W(T)} = \mathbb{P}\left(\xi_T\left(\widetilde{\mathcal{T}}_\top\right) \neq 0\right) = \widetilde{P}_{\mu_\top}\left(\tau_T^+ < \tau_0\right) = P_T\left(\tau_0 < \tau_T^+\right) \\ \widetilde{\gamma} &= \frac{1}{\widetilde{W}(T)} = \mathbb{P}\left(\xi_T\left(\mathcal{T}_\perp\right) \neq 0\right) = P_{\mu_\perp}\left(\tau_T^+ < \tau_0\right) = \widetilde{P}_T\left(\tau_0 < \tau_T^+\right) \end{aligned}$$

Contour of a forest

Y Lévy process with measure Π and drift -1, reflected at T and stopped at

$$\tau := \inf\{t \geq \tau_T : Y_t^{(T)} = 0\}$$

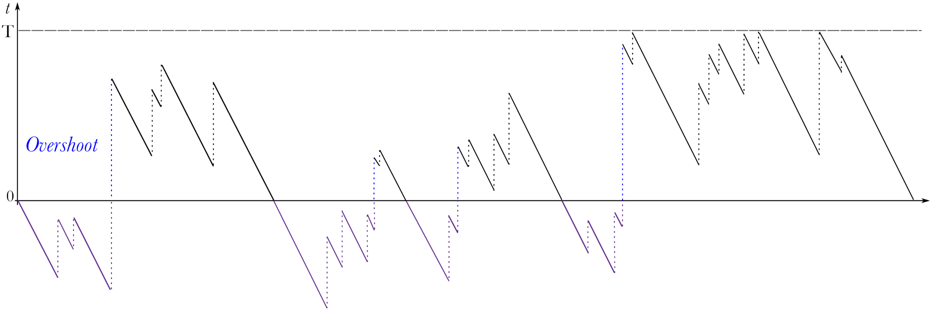


Contour of a forest

Define the time change

$$A_t := \int_0^t \mathbb{1}_{\{Y_u > 0\}} \mathrm{d}u$$

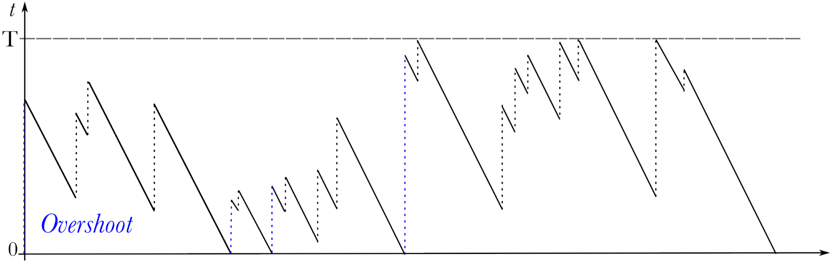
and its right-continuous inverse, $\alpha(t) := \inf\{u \geq 0 : A_u > t\}$.



Contour of a forest

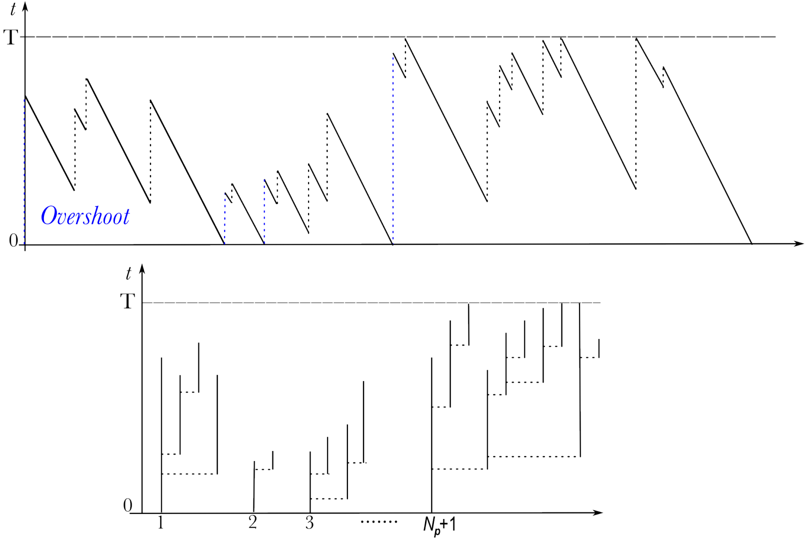
Consider

$$Y^{(T)} \circ \alpha$$



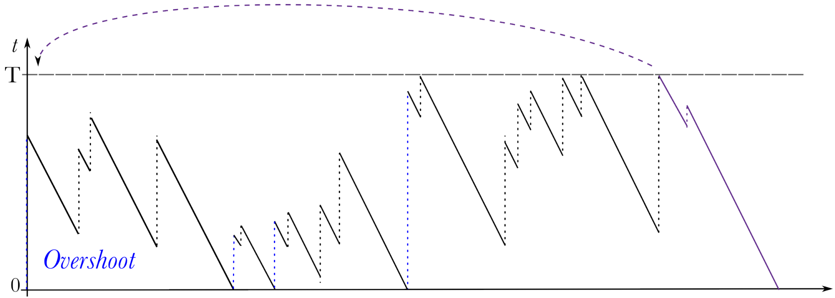
Contour of a forest

Then we have $Y^{(T)} \circ \alpha \overset{d}{=} \mathcal{C}\left(\mathcal{F}_{\perp}^{\tilde{\gamma}}\right)$



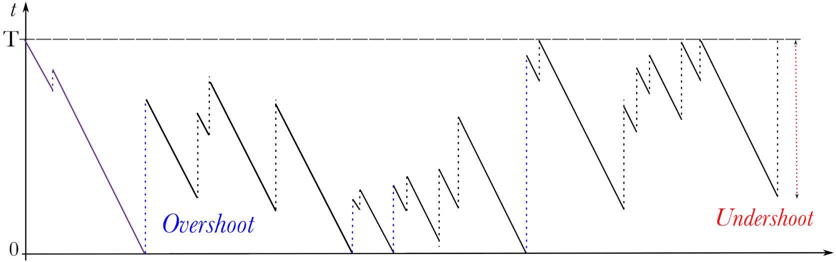
Contour of a forest

Now consider $\chi\left(Y^{(T)} \circ \alpha\right)$



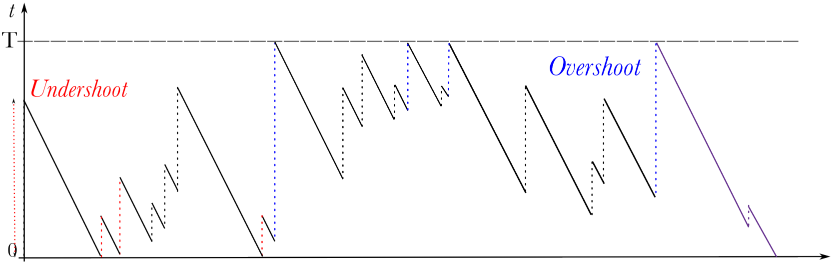
Contour of a forest

Now consider $\chi\left(Y^{(T)} \circ \alpha\right)$



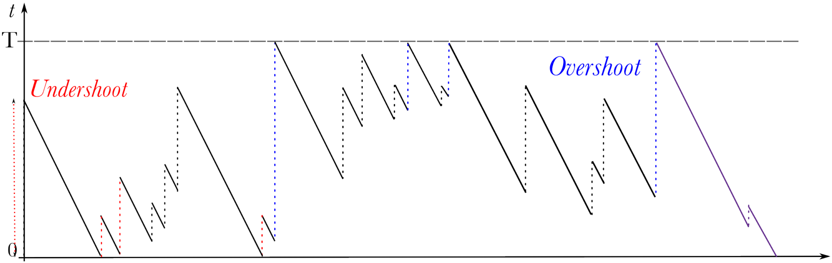
Contour of a forest

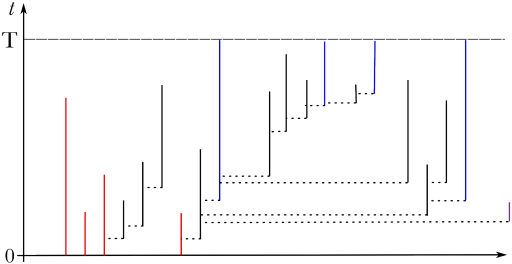
And the space-time-reversed process $\rho \circ \chi\left(Y^{(T)} \circ \alpha\right)$



Contour of a forest

And we have again $\rho \circ \chi \left(Y^{(T)} \circ \alpha \right) \stackrel{d}{=} \mathcal{C}\left(\widetilde{\mathcal{F}}^{\gamma}_{\top} \right)$





Outline

1. Birth-death models
2. Time reversal for birth-death processes
3. Application to epidemiology
4. Generalization for non-exponential infectiousness periods
5. Ingredients of the proof
6. Conclusions, perspectives and some insight

Conclusions and perspectives

Conclusions

- We have obtained a duality under time-reversal for BD and general branching forests (Crump-Mode-Jagers's processes), stopped at their first surviving tree. The duality concerns the population size processes and the genealogies.
- Applications to phylodynamic modeling:
  - the results allow to characterize the population size process distribution conditionally on the reconstructed transmission tree;
  - access to the likelihood to allow inference using prevalence time series and sequence data

Ongoing work

Consequent results on epidemics seen from their end. In particular, invariance by time reversal of the excursion away from 0 of the critical Feller diffusion, which is the width process of the continuum random tree.

Traveling backwards in time: insight on duality

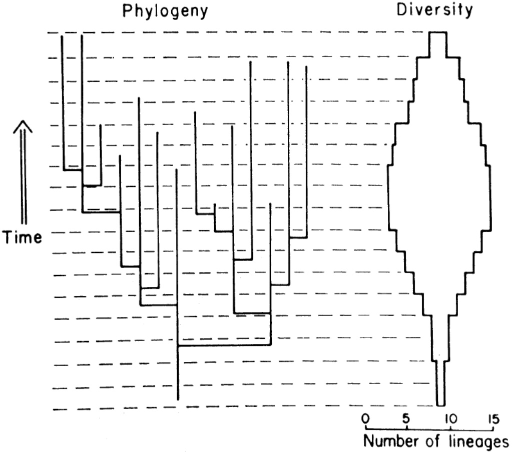


Fig. 1.—Hypothetical phylogeny (*left*) and the pattern of change in number of lineages through time (*right*).

Figures from [Raup et al., 1973]

Traveling backwards in time: insight on duality

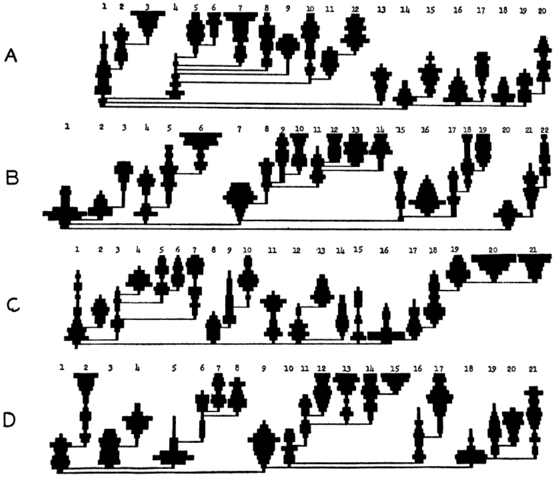


FIG. 5.—Diversity variation and phylogeny of the clades produced by four computer runs. Run *A* is the same as that used for figs. 3 and 4. Clades *A*-19 and *A*-20 are shown in detail in fig. 3. Runs *B*, *C*, and *D* used the same input constants but different random number sequences.

Figures from [Raup et al., 1973]

References I

J. Bertoin
An Extension of Pitman's Theorem for Spectrally Positive Levy Processes
Ann. Probab., 20(3):1464–1483,1992.

J. Bertoin
Lévy Processes
Cambridge University Press, Cambridge, Vol 191, 1996.

M. Dávila Felipe and A. Lambert
Time reversal dualities for some random forests
ALEA, Lat. Am. J. Probab. Math. Stat. 12 (1), 399-426, 2015.

B.T. Grenfell, O.G. Pybus, J.R. Gog, J.L Wood, J.M. Daly, J.A. Mumford and E.C. Holmes
Unifying the Epidemiological and Evolutionary Dynamics of Pathogens
Science, 303: 327-332, 2004.

A. Lambert
The contour of splitting trees is a Lévy process
Ann. Probab., 38(1):348–395, 2010.

References II

A. Lambert, H.K. Alexander and T. Stadler
Phylogenetic analysis accounting for age-dependent death and sampling with applications to epidemics
Journal of Theoretical Biology, 352(7):60–70, 2014.

D.A. Rasmussen, O. Ratman, and K. Koelle
Inference for nonlinear epidemiological models using genealogies and time series
PLoS Comput. Biol., 7(8):e1002136, 2011.

D.M. Raup, S.J. Gould, T.J.M. Schopf and D.S. Simberloff
Stochastic Models of Phylogeny and the Evolution of Diversity
The Journal of Geology, 81(5): 525–542, 1973.

S.V. Scarpino
Evolutionary Medicine IV. Evolution and Emergence of Novel Pathogens
Encyclopedia of Evolutionary Biology, 2: 77-?82, 2016.

T. Stadler et al.
Estimating the Basic Reproductive Number from Viral Sequence Data
Molecular Biology and Evolution, 29(1), 347–357, 2012.

References III

E. Volz, K. Koelle, and T. Bedford
Viral Phylodynamics
PLoS Comput Biol, 9(3): e1002947, 2013.

E. Volz, S. Pond, M. Ward, A. Leigh Brown, and S. Frost
Phylodynamics of infectious disease epidemics.
Genetics, 183(4):1421–1430, 2009.

Thanks!


π
S.M.I.L.E
Stochastic Models for the Inference of Life Evolution