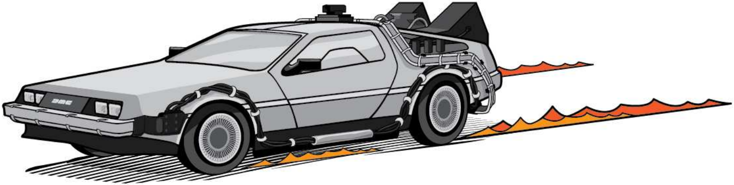


A Computational Time Machine

with possible applications to phylogenetic tree reconstruction

Tim Vaughan*, Alexei Drummond†, Peter Drummond*

*Swinburne University of Technology
Melbourne, Australia

†The University of Auckland,
Auckland, New Zealand

5th of November, 2009

Overview of Talk

Overview of Talk

Part 1 A Systematic Method for Retrodictive Inference

Overview of Talk

Part 1 A Systematic Method for Retrodictive Inference

- Motivation and theoretical context

Overview of Talk

Part 1 A Systematic Method for Retrodictive Inference

- Motivation and theoretical context
- ‘Retrodictive’ Master Equations

Overview of Talk

Part 1 A Systematic Method for Retrodictive Inference

- Motivation and theoretical context
- ‘Retrodictive’ Master Equations
- Stochastic solution of RMEs (Retrodictive SSA)

Overview of Talk

Part 1 A Systematic Method for Retrodictive Inference

- Motivation and theoretical context
- ‘Retrodictive’ Master Equations
- Stochastic solution of RMEs (Retrodictive SSA)

Part 2 Application of the RSSA to Phylogenetic Tree Reconstruction

Part I

Systematic methods for Retrodictive Inference

Continuous-time birth/death processes

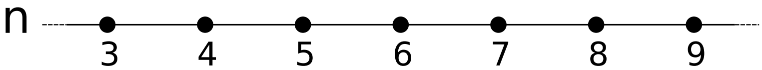
- Govern the markovian stochastic dynamics of systems which undergo discrete state changes at random times.

Continuous-time birth/death processes

- Govern the markovian stochastic dynamics of systems which undergo discrete state changes at random times.
- Employed to describe a vast array of physical processes in
  - physics (statistical and quantum mechanics),
  - chemistry (chemical reactions) and
  - biology (stochastic population dynamics, sequence evolution models, etc. etc.).

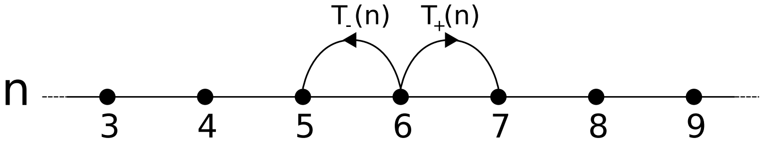
Continuous-time birth/death processes

A single-variable example:



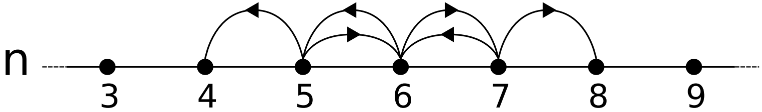
Continuous-time birth/death processes

A single-variable example:



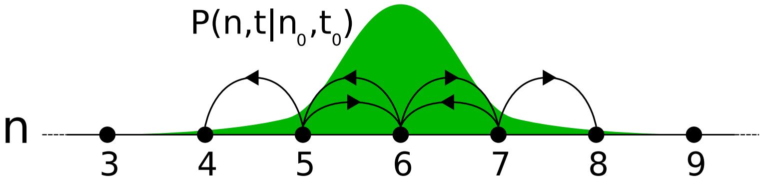
Continuous-time birth/death processes

A single-variable example:



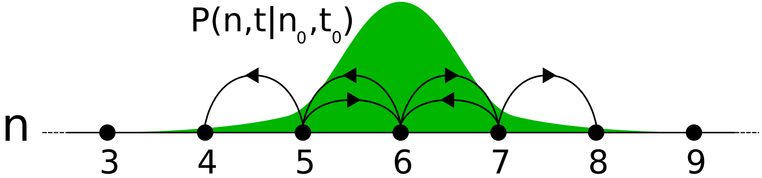
Continuous-time birth/death processes

A single-variable example:



Continuous-time birth/death processes

A single-variable example:



In general, probability distributions of continuous-time birth/death processes satisfy the

Master Equation

$$\frac{d}{dt}P(\vec{n},t|\vec{n}_0,t_0) = \sum_k \left[T_k(\vec{n}'_k)P(\vec{n}'_k,t) - T_k(\vec{n})P(\vec{n},t) \right]$$

Looking into the past: retrodictive inference

Looking into the past: retrodictive inference

- As solution of the ME yields conditional probabilities of later states given earlier states, it is naturally suited to predictive inference.

Looking into the past: retrodictive inference

- As solution of the ME yields conditional probabilities of later states given earlier states, it is naturally suited to predictive inference.
- To use $P(\vec{n}_f, t_f | \vec{n}_i, t_i)$ to infer earlier states from later states (i.e. retrodiction), employ Bayes' theorem:

Looking into the past: retrodictive inference

- As solution of the ME yields conditional probabilities of later states given earlier states, it is naturally suited to predictive inference.
- To use $P(\vec{n}_f, t_f | \vec{n}_i, t_i)$ to infer earlier states from later states (i.e. retrodiction), employ Bayes' theorem:

$$P(\vec{n}_i, t_i | \vec{n}_f, t_f) = \frac{P(\vec{n}_f, t_f | \vec{n}_i, t_i) P_0(\vec{n}_i, t_i)}{\sum_{\vec{m}} P(\vec{n}_f, t_f | \vec{m}, t_i) P_0(\vec{m}, t_i)}$$

Looking into the past: retrodictive inference

- As solution of the ME yields conditional probabilities of later states given earlier states, it is naturally suited to predictive inference.
- To use $P(\vec{n}_f, t_f | \vec{n}_i, t_i)$ to infer earlier states from later states (i.e. retrodiction), employ Bayes' theorem:

$$P(\vec{n}_i, t_i | \vec{n}_f, t_f) = \frac{P(\vec{n}_f, t_f | \vec{n}_i, t_i) P_0(\vec{n}_i, t_i)}{\sum_{\vec{m}} P(\vec{n}_f, t_f | \vec{m}, t_i) P_0(\vec{m}, t_i)}$$

- Requires either MCMC or solving a prohibitively large number of initial value problems.

Looking into the past: retrodictive inference

- As solution of the ME yields conditional probabilities of later states given earlier states, it is naturally suited to predictive inference.
- To use $P(\vec{n}_f, t_f | \vec{n}_i, t_i)$ to infer earlier states from later states (i.e. retrodiction), employ Bayes' theorem:

$$P(\vec{n}_i, t_i | \vec{n}_f, t_f) = \frac{P(\vec{n}_f, t_f | \vec{n}_i, t_i) P_0(\vec{n}_i, t_i)}{\sum_{\vec{m}} P(\vec{n}_f, t_f | \vec{m}, t_i) P_0(\vec{m}, t_i)}$$

- Requires either MCMC or solving a prohibitively large number of initial value problems.

Is there a more direct approach?

Looking into the past: retrodictive inference

A completely equivalent description of the stochastic dynamics is provided by the . . .

Looking into the past: retrodictive inference

A completely equivalent description of the stochastic dynamics is provided by the ...

Backward Master Equation

$$\begin{aligned}\frac{d}{dt_0}P(\vec{n},t|\vec{n}_0,t_0) &= \sum_k T_k(\vec{n}_0)\big[P(\vec{n},t|\vec{n}''_{0k},t_0) \\ &\quad -P(\vec{n},t|\vec{n}_0,t_0)\big]\end{aligned}$$

Looking into the past: retrodictive inference

A completely equivalent description of the stochastic dynamics is provided by the ...

Backward Master Equation

$$\begin{aligned}\frac{d}{dt_0}P(\vec{n},t|\vec{n}_0,t_0) &= \sum_k T_k(\vec{n}_0)\big[P(\vec{n},t|\vec{n}''_{0k},t_0)\\ &\quad -P(\vec{n},t|\vec{n}_0,t_0)\big]\end{aligned}$$

(Can be derived from the regular ME by noting that $d/ds\sum_{\vec{m}} P(\vec{n},t|\vec{m},s)P(\vec{m},s|\vec{n}_0,t_0) = 0$ for $t > s > t_0$.)

Looking into the past: retrodictive inference

- For convenience, define

$$R(\vec{n}, \tau | \vec{n}_f) = \frac{P(\vec{n}_f, t_f | \vec{n}, t_f - \tau)}{\sum_{\vec{m}} P(\vec{n}_f, t_f | \vec{m}, t_f - \tau)}$$

which is equivalent to the posterior distribution over states at the earlier time $t_f - \tau$, given a uniform prior.

Looking into the past: retrodictive inference

- For convenience, define

$$R(\vec{n},\tau|\vec{n}_f) = \frac{P(\vec{n}_f,t_f|\vec{n},t_f-\tau)}{\sum_{\vec{m}} P(\vec{n}_f,t_f|\vec{m},t_f-\tau)}$$

which is equivalent to the posterior distribution over states at the earlier time $t_f-\tau$, <u>given a uniform prior</u>.

- Via the BME one finds $R(\vec{n},\tau|\vec{n}_f)$ evolves due to the ...

Retrodictive Master Equation

$$\begin{aligned}\frac{d}{d\tau}R(\vec{n},\tau|\vec{n}_f) &= \sum_k T_k(\vec{n})\big[R(\vec{n}''_k,\tau|\vec{n}_f) - R(\vec{n},\tau|\vec{n}_f)\big] \\ &\quad -\dot{\Omega}(\tau)R(\vec{n},\tau|\vec{n}_f)\end{aligned}$$

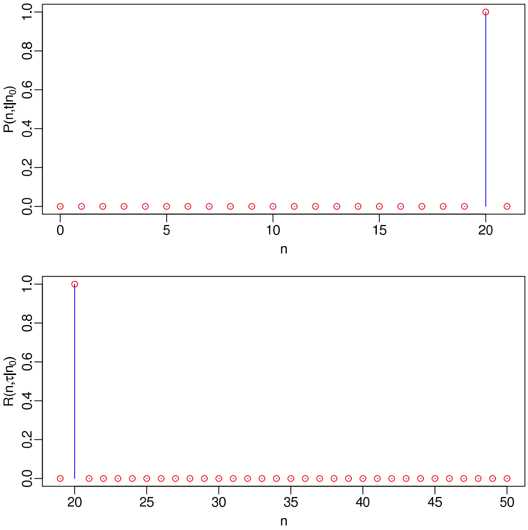
where $\dot{\Omega}(\tau)$ preserves global normalisation.

Barnett, Pegg, Jeffers & Jedrkiewicz, PRL, 2001

Example: pure death process

- Consider process governed by the reaction:

$$X \longrightarrow 0$$



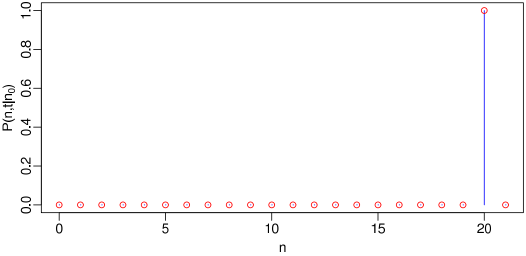
Example: pure death process

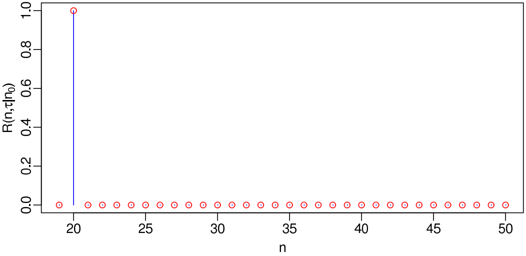
- Consider process governed by the reaction:

$$X \longrightarrow 0$$

- Figures show $P(n,t)$ and $R(n,\tau)$ at

$$t = \tau = 0$$





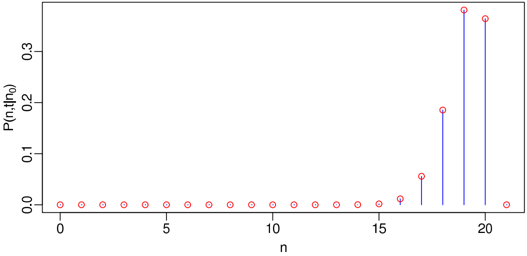
Example: pure death process

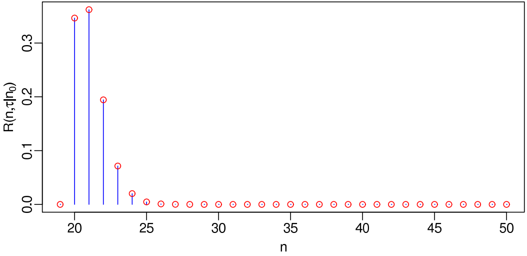
- Consider process governed by the reaction:

$$X \longrightarrow 0$$

- Figures show $P(n,t)$ and $R(n,\tau)$ at

$$t = \tau = 0.05$$





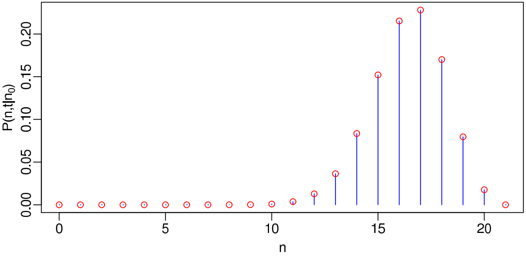
Example: pure death process

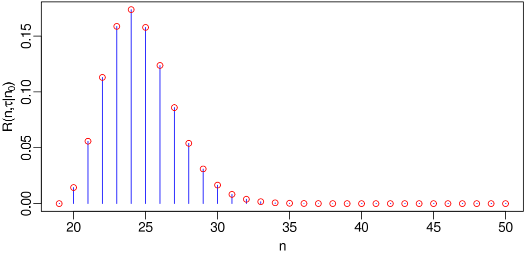
- Consider process governed by the reaction:

$$X \longrightarrow 0$$

- Figures show $P(n,t)$ and $R(n,\tau)$ at

$$t = \tau = 0.2$$





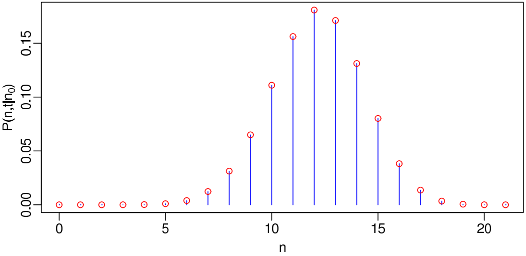
Example: pure death process

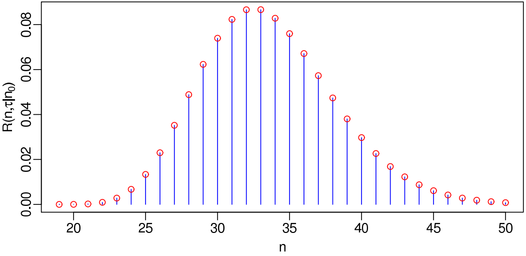
- Consider process governed by the reaction:

  $$X \longrightarrow 0$$

- Figures show $P(n,t)$ and $R(n,\tau)$ at

  $$t = \tau = 0.5$$





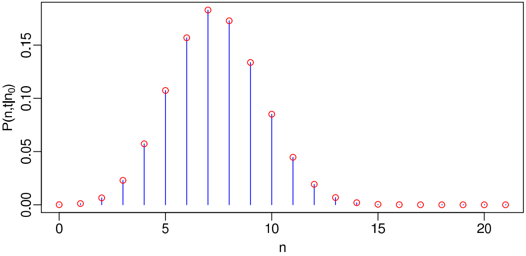
Example: pure death process

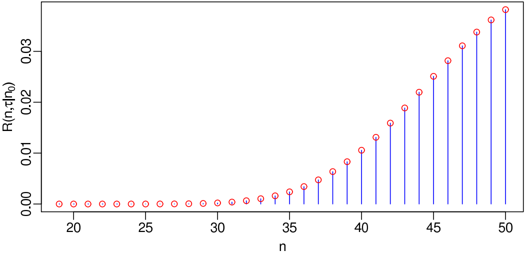
- Consider process governed by the reaction:

$$X \longrightarrow 0$$

- Figures show $P(n,t)$ and $R(n,\tau)$ at

$$t = \tau = 1.0$$

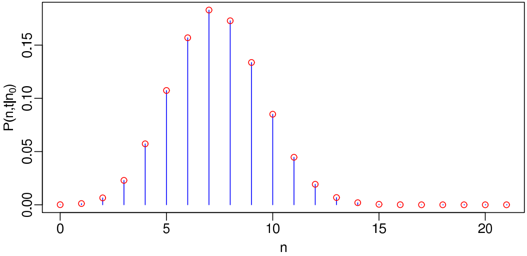




Example: pure death process

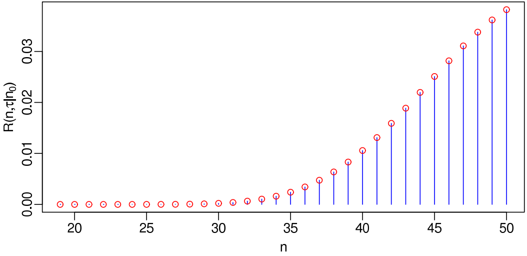
- Consider process governed by the reaction:

$$X \longrightarrow 0$$



- Figures show $P(n,t)$ and $R(n,\tau)$ at

$$t = \tau = 1.0$$



Note: $R(\vec{n},\tau)$ is <u>not</u> simply time-reversed $P(\vec{n},t)$!

The problem of complexity

The problem of complexity

- Direct integration of master equations (either predictive or retrodictive) is usually impossible in practice.

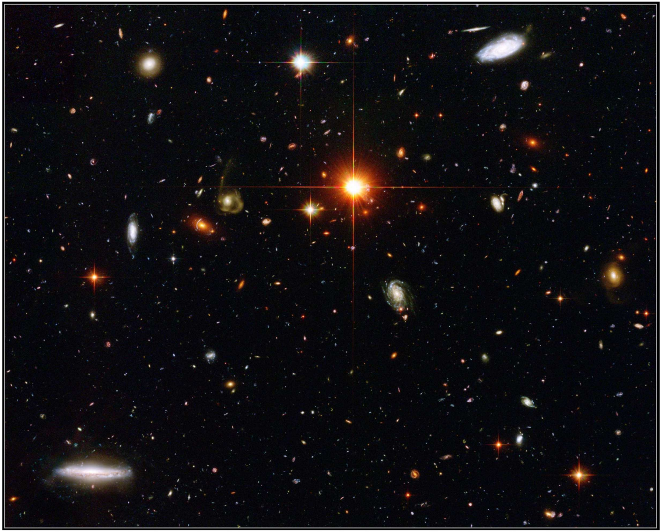
The problem of complexity

- Direct integration of master equations (either predictive or retrodictive) is usually impossible in practice.
- This is due to the vast state space available to the dynamics of even simple systems.

The problem of complexity

- Direct integration of master equations (either predictive or retrodictive) is usually impossible in practice.
- This is due to the vast state space available to the dynamics of even simple systems.
- For example, a population composed of 30 types of organism, with each subpopulation having < 100 members can exist in any one of over $(100^2/2)^{30} \sim 10^{110}$ distinct states.

The problem of complexity



NASA, ESA, and The Hubble Heritage Team (STScI/AURA) • Hubble Space Telescope ACS • STScI-PRC05-20 Hubble Heritage

There are approximately 10^{80} atoms in the visible universe.

Solution: Stochastic Simulation Algorithms (SSAs)

Solution: Stochastic Simulation Algorithms (SSAs)

Central Idea

Sacrifice *precision* for *computability*.

Solution: Stochastic Simulation Algorithms (SSAs)

Central Idea

Sacrifice *precision* for *computability*.

- Generally credited to the work of Daniel T. Gillespie.

J. Comp. Phys. (1976)

Solution: Stochastic Simulation Algorithms (SSAs)

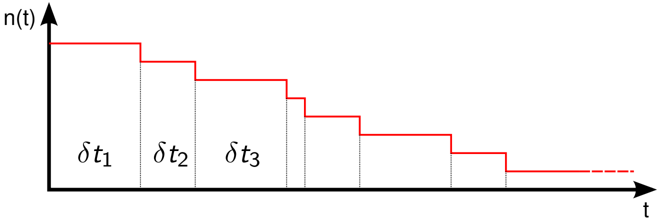
Central Idea

Sacrifice *precision* for *computability*.

- Generally credited to the work of Daniel T. Gillespie.

J. Comp. Phys. (1976)

- For forward-time MEs, involve literally simulating the birth/death process.



Solution: Stochastic Simulation Algorithms (SSAs)

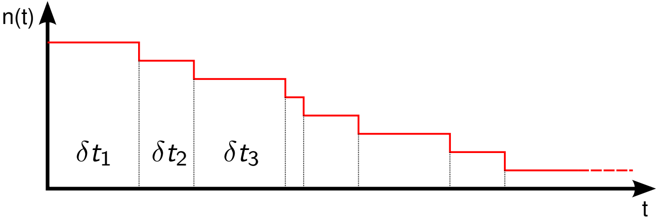
Central Idea

Sacrifice *precision* for *computability*.

- Generally credited to the work of Daniel T. Gillespie.

J. Comp. Phys. (1976)

- For forward-time MEs, involve literally simulating the birth/death process.



- The simulated trajectories can be considered picks from the solution to the associated ME.

Solution: Stochastic Simulation Algorithms (SSAs)

- An equivalent expression the dynamics described by ME is given by the following “reaction probability density function”:

$$p(k, \delta t | \vec{n}, t) = T_k(\vec{n}) \exp[-\delta t \sum_k T_k(\vec{n})]$$

Solution: Stochastic Simulation Algorithms (SSAs)

- An equivalent expression the dynamics described by ME is given by the following “reaction probability density function”:

$$p(k, \delta t | \vec{n}, t) = T_k(\vec{n}) \exp[-\delta t \sum_k T_k(\vec{n})]$$

Algorithm

1. Initialize trajectory state to $\vec{n}_0$ or $P(\vec{n}_0)$,
2. Draw δt and reaction type k from RPDF,
3. Update trajectory state and time,
4. If trajectory time $<$ finishing time, GOTO 2

Solution: Stochastic Simulation Algorithms (SSAs)

- An equivalent expression the dynamics described by ME is given by the following "reaction probability density function":

$$p(k, \delta t | \vec{n}, t) = T_k(\vec{n}) \exp[-\delta t \sum_k T_k(\vec{n})]$$

Algorithm

1. Initialize trajectory state to $\vec{n}_0$ or $P(\vec{n}_0)$,
2. Draw δt and reaction type k from RPDF,
3. Update trajectory state and time,
4. If trajectory time $<$ finishing time, GOTO 2

- The relative frequency of trajectories posessing state $\vec{n}$ at time t converges to $P(\vec{n}, t | \vec{n}_0, t_0)$.

A Stochastic Interpretation of RMEs

A Stochastic Interpretation of RMEs

There are intrinsic difficulties in finding stochastic solutions to retrodictive master equations:

A Stochastic Interpretation of RMEs

There are intrinsic difficulties in finding stochastic solutions to retrodictive master equations:

- Equation of motion is nonlinear due to the normalisation term.

$$\begin{aligned}\frac{\partial}{\partial\tau}R(\vec{n},\tau|\vec{n}_f) &= \sum_k T_k(\vec{n})\left[R(\vec{n}''_k,\tau|\vec{n}_f) - R(\vec{n},\tau|\vec{n}_f)\right] \\ &\quad -\dot{\Omega}(\tau)R(\vec{n},\tau|\vec{n}_f)\end{aligned}$$

A Stochastic Interpretation of RMEs

There are intrinsic difficulties in finding stochastic solutions to retrodictive master equations:

- Equation of motion is nonlinear due to the normalisation term.

$$\begin{aligned}\frac{\partial}{\partial \tau} R(\vec{n}, \tau | \vec{n}_f) &= \sum_k T_k(\vec{n}) \left[R(\vec{n}''_k, \tau | \vec{n}_f) - R(\vec{n}, \tau | \vec{n}_f) \right] \\ &\quad - \dot{\Omega}(\tau) R(\vec{n}, \tau | \vec{n}_f)\end{aligned}$$

- Remaining linear component does not preserve normalization, meaning that there is no description in terms of relative frequencies of single unweighted trajectory states.

A Stochastic Interpretation of RMEs

There are intrinsic difficulties in finding stochastic solutions to retrodictive master equations:

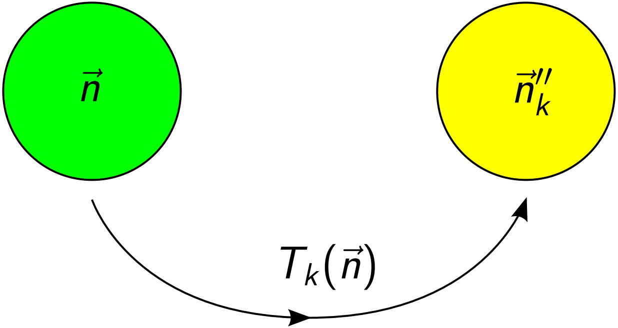
- Equation of motion is nonlinear due to the normalisation term.

$$\begin{aligned}\frac{\partial}{\partial\tau}R(\vec{n},\tau|\vec{n}_f) &= \sum_k T_k(\vec{n})\left[R(\vec{n}''_k,\tau|\vec{n}_f) - R(\vec{n},\tau|\vec{n}_f)\right] \\ &\quad -\dot{\Omega}(\tau)R(\vec{n},\tau|\vec{n}_f)\end{aligned}$$

- Remaining linear component does not preserve normalization, meaning that there is no description in terms of relative frequencies of single unweighted trajectory states.

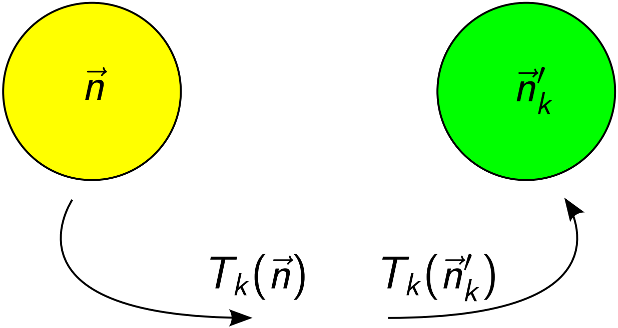
These difficulties are related to the fact that $R(\vec{n},\tau|\vec{n}_f)$ describes incomplete information – *not* the relative frequencies of outcomes of a random experiment.

A Stochastic Interpretation of RMEs



Predictive master equation transition rate

A Stochastic Interpretation of RMEs



$\vec{n}$

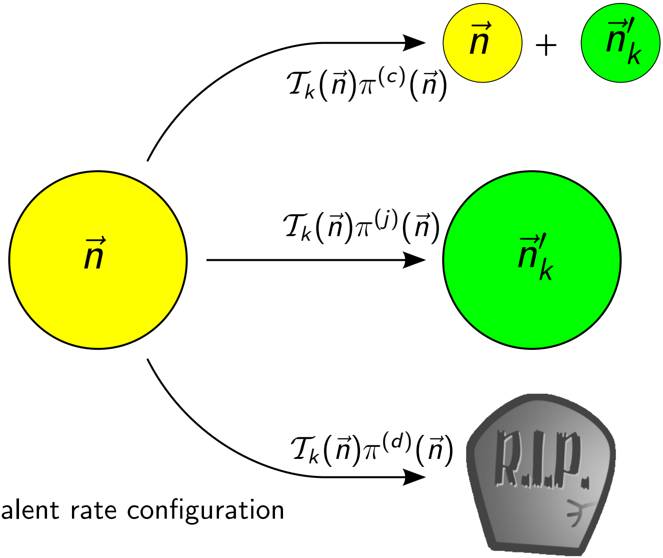
$\vec{n}'_k$

$T_k(\vec{n})$

$T_k(\vec{n}'_k)$

Retrodictive master equation (un-normalised) transition rates

A Stochastic Interpretation of RMEs



$\vec{n}$ + $\vec{n}'_k$

$\mathcal{T}_k(\vec{n})\pi^{(c)}(\vec{n})$

$\vec{n}$

$\mathcal{T}_k(\vec{n})\pi^{(j)}(\vec{n})$

$\vec{n}'_k$

$\mathcal{T}_k(\vec{n})\pi^{(d)}(\vec{n})$

Equivalent rate configuration

A Stochastic Interpretation of RMEs

Here we have defined a composite reaction which occurs at rate:

$$\mathcal{T}_k(\vec{n}) = \max[T_k(\vec{n}), T_k(\vec{n}'_k)]$$

and involves either a standard jump, the creation of a new trajectory, or the deletion of the current trajectory with the respective probabilities:

$$\begin{aligned}
\pi_k^{(j)}(\vec{n}) &= \min[T_k(\vec{n}), T_k(\vec{n}'_k)], \\
\pi_k^{(c)}(\vec{n}) &= \max[T_k(\vec{n}'_k) - T_k(\vec{n}'_k), 0] \text{ and} \\
\pi_k^{(d)}(\vec{n}) &= \max[T_k(\vec{n}) - T_k(\vec{n}'_k), 0].
\end{aligned}$$

A Stochastic Interpretation of RMEs

- Can write the RME explicitly in terms of these non-conservative reactions:

$$\begin{aligned}\frac{\partial}{\partial \tau} R(\vec{n}, \tau | \vec{n}_f) &= \sum_k \Big\{ \mathcal{T}_k(\vec{n}''_k) \left[\pi_k^{(j)}(\vec{n}''_k) + \pi_k^{(c)}(\vec{n}''_k) \right] R(\vec{n}''_k, \tau | \vec{n}_f) \\ &\quad - \mathcal{T}_k(\vec{n}) \left[\pi_k^{(j)}(\vec{n}) + \pi_k^{(d)}(\vec{n}) \right] R(\vec{n}, \tau | \vec{n}_f) \Big\} \\ &\quad - \dot{\Omega}(\tau) R(\vec{n}, \tau | \vec{n}_f)\end{aligned}$$

A Stochastic Interpretation of RMEs

- Thus, just as for the predictive master equation, an equivalent expression of the dynamics described by the RME is given by the modified reaction probability density function:

$$p(k, \delta\tau | \vec{n}) = \mathcal{T}_k(\vec{n}) \exp[-\delta\tau \sum_k \mathcal{T}_k(\vec{n})]$$

A Stochastic Interpretation of RMEs

- Thus, just as for the predictive master equation, an equivalent expression of the dynamics described by the RME is given by the modified reaction probability density function:

$$p(k, \delta\tau|\vec{n}) = \mathcal{T}_k(\vec{n}) \exp[-\delta\tau \sum_k \mathcal{T}_k(\vec{n})]$$

General Retrodictive SSA

1. Initialise trajectory state to $\vec{n}_f$ or $P(\vec{n}_f)$,
2. draw $\delta\tau$ and k from modified RPDF,
3. Select a jump, creation or deletion with probabilities $\pi_k^{(j)}$,$\pi_k^{(c)}$ and $\pi_k^{(d)}$, respectively.
4. Update trajectory state and time according to chosen reaction.
5. Fork process to follow created trajectory, terminate process of deleted trajectory.
6. If trajectory time $<$ finishing time, GOTO 2.

Practical Implementation Issues

Practical Implementation Issues

Clearly, literal implementation of this algorithm may result either

Practical Implementation Issues

Clearly, literal implementation of this algorithm may result either

1. unbounded growth in the number of trajectories, or

Practical Implementation Issues

Clearly, literal implementation of this algorithm may result either

1. unbounded growth in the number of trajectories, or
2. premature depletion of the trajectory ensemble.

Practical Implementation Issues

Clearly, literal implementation of this algorithm may result either

1. unbounded growth in the number of trajectories, or
2. premature depletion of the trajectory ensemble.

Solution

Explicitly or implicitly resample the trajectory ensemble.

Practical Implementation Issues

Resampling will eventually cause the trajectories in the ensemble to become artificially correlated, leading to systematic errors. The magnitude of these errors will vary inversely with the size of the ensemble.

Practical Implementation Issues

Resampling will eventually cause the trajectories in the ensemble to become artificially correlated, leading to systematic errors. The magnitude of these errors will vary inversely with the size of the ensemble.

Solution

Repeat each calculation with progressively larger trajectory ensembles, until the result converges to a fixed value.

Practical Implementation Issues

Realistic models of population genetics can involve a large number of reaction types M, and the computational complexity of naïve implementations of the reaction-selection scale as $O(M)$.

Practical Implementation Issues

Realistic models of population genetics can involve a large number of reaction types M, and the computational complexity of naive implementations of the reaction-selection scale as $O(M)$.

Solution

Just as for the Gibson-Bruck modification to the predictive SSA, use binary search trees to change the scaling to $O(\log(M))$.

Simple example: The Moran model

Simple example: The Moran model

Consider a fixed population of 2 alleles (a and A) evolving under a neutral continuous-time Moran model:

$$\begin{aligned} a + A &\longrightarrow 2A \\ a + A &\longrightarrow 2a \end{aligned}$$

Simple example: The Moran model

Consider a fixed population of 2 alleles (a and A) evolving under a neutral continuous-time Moran model:

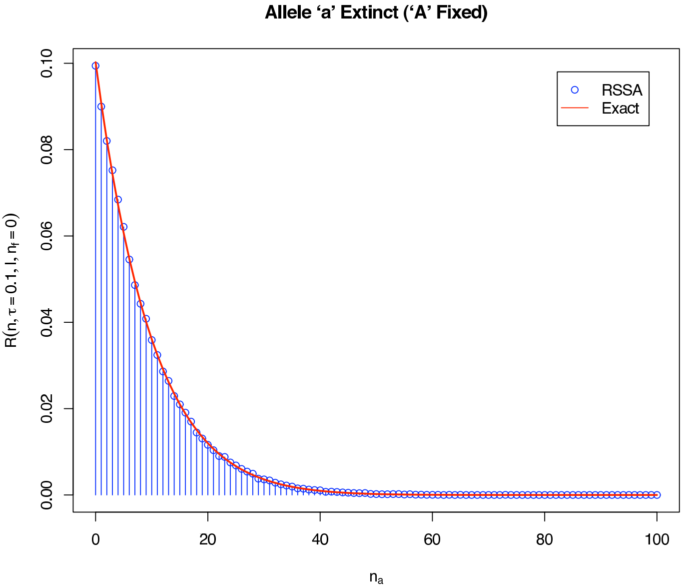
$$
\begin{aligned}
a + A &\longrightarrow 2A \\
a + A &\longrightarrow 2a
\end{aligned}
$$

The RME for this system is

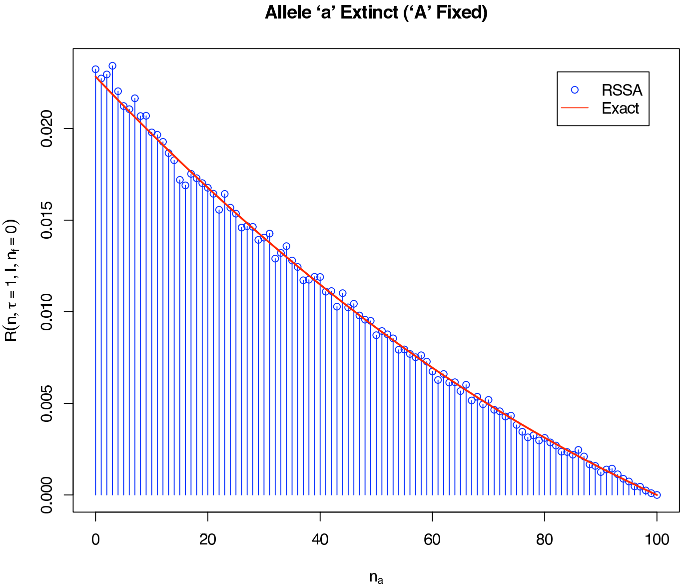
$$
\begin{aligned}
\frac{\partial}{\partial \tau} R(n, \tau | n_f) &= n(N-n)\Big[R(n+1, \tau | n_f) + R(n-1, \tau | n_f) \\
&\quad -2R(n, \tau | n_f)\Big] - \dot{\Omega}(\tau) R(n, \tau | n_f)
\end{aligned}
$$

where n is the number of individuals carrying the a allele and N is the total population size.

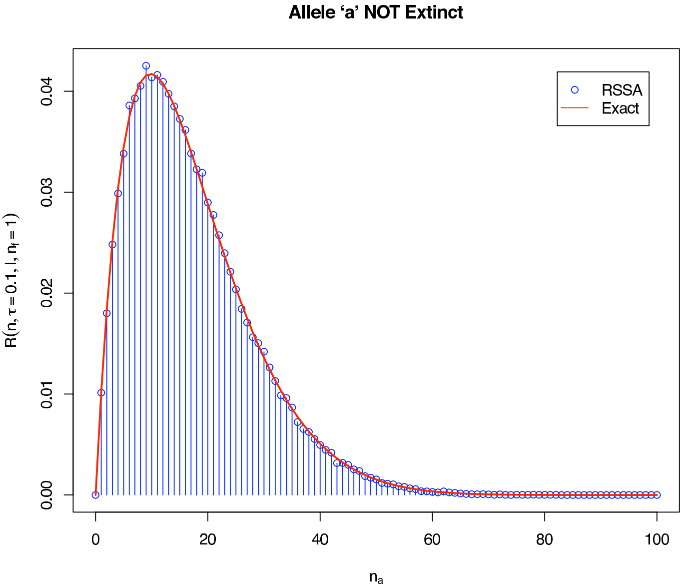
Simple example: The Moran model



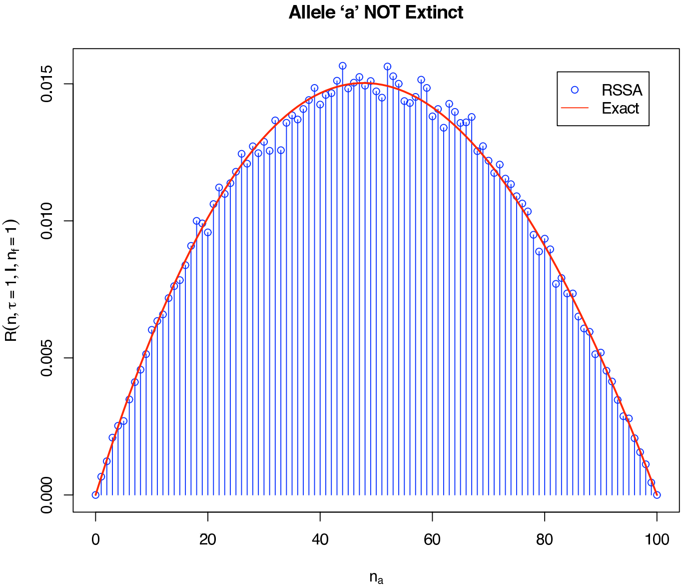
Simple example: The Moran model



Simple example: The Moran model



Simple example: The Moran model



Possible Enhancements

Possible Enhancements

- Employ weighted trajectories to provide stochastic solutions to the RME rather than trajectory ensembles of varying size. This would remove primary reason for resampling and therefore do away with the systematic error buildup this causes.

Possible Enhancements

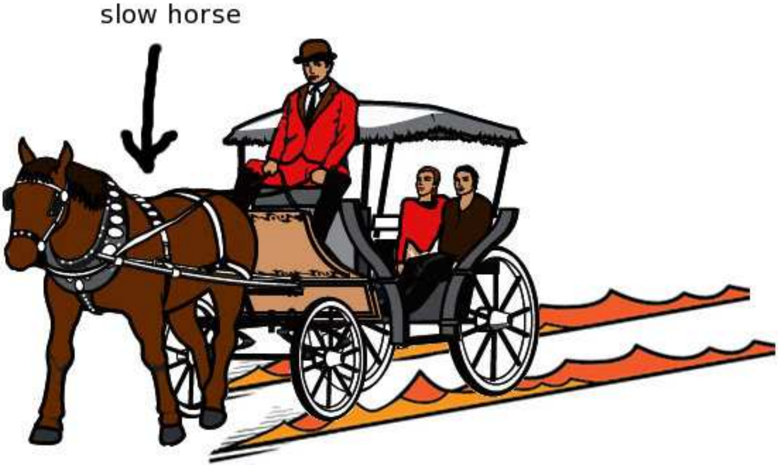
- Employ weighted trajectories to provide stochastic solutions to the RME rather than trajectory ensembles of varying size. This would remove primary reason for resampling and therefore do away with the systematic error buildup this causes.
- Follow the hybrid trajectory branching / weighting approach which is used in Quantum Monte Carlo calculations of ground states of many-body systems, which are reported to reduce fluctuations in trajectory weights and therefore improve the precision of estimated moments.

Possible Enhancements

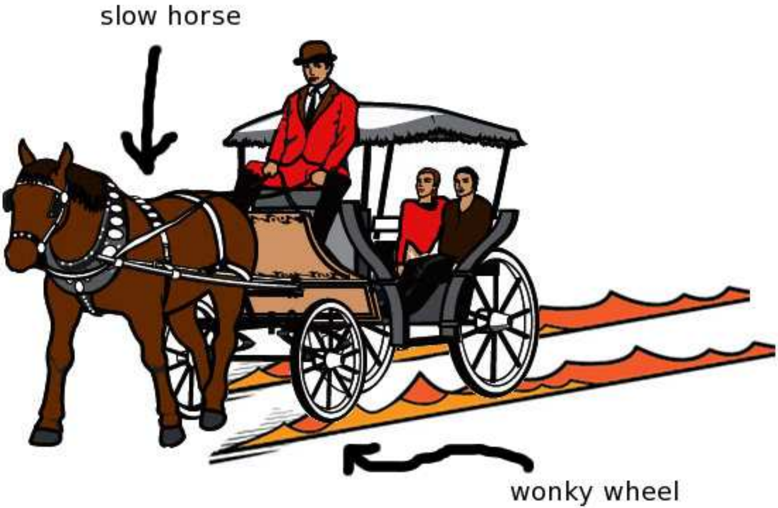
- Employ weighted trajectories to provide stochastic solutions to the RME rather than trajectory ensembles of varying size. This would remove primary reason for resampling and therefore do away with the systematic error buildup this causes.
- Follow the hybrid trajectory branching / weighting approach which is used in Quantum Monte Carlo calculations of ground states of many-body systems, which are reported to reduce fluctuations in trajectory weights and therefore improve the precision of estimated moments.
- Investigate whether phase-space methods can be used to dramatically improve the scalability of these calculations to large populations.

Possible Enhancements

Possible Enhancements



Possible Enhancements



Part II

Applications to Phylogenetic Tree Reconstruction

Possible uses for RSSAs in Phylogenetics

Possible uses for RSSAs in Phylogenetics

1. Full retrodictive inference of the past system history given a realistic stochastic model of evolutionary dynamics.

Possible uses for RSSAs in Phylogenetics

1. Full retrodictive inference of the past system history given a realistic stochastic model of evolutionary dynamics.
2. A systematic means of combining forwards-time mutation/migration dynamics with the Kingman *n*-coalescent process.

Full retrodictive inference

Full retrodictive inference

Bad News

This is computationally intractable using SSA-style algorithms for most systems of realistic size.

Full retrodictive inference

Bad News

This is computationally intractable using SSA-style algorithms for most systems of realistic size.

Potentially tractable problems include:

Full retrodictive inference

Bad News

This is computationally intractable using SSA-style algorithms for most systems of realistic size.

Potentially tractable problems include:

- Exponential growth stage of within-host viral infection.

Full retrodictive inference

Bad News

This is computationally intractable using SSA-style algorithms for most systems of realistic size.

Potentially tractable problems include:

- Exponential growth stage of within-host viral infection.
- Spatially constricted populations with small carrying capacities.

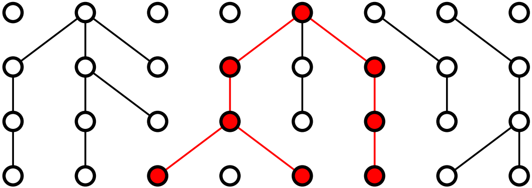
Combining forward processes with the Coalescent

Good News

Coalescent formalism allows us to avoid explicit consideration of the bulk of a population by focusing on sampled individuals.

Combining forward processes with the Coalescent

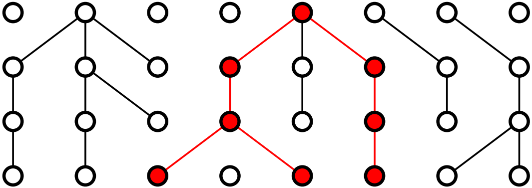
Start with basic coalescent:



Coalescence rate $T_C(k) \simeq \frac{1}{2N} k(\tau)(k(\tau) - 1)$

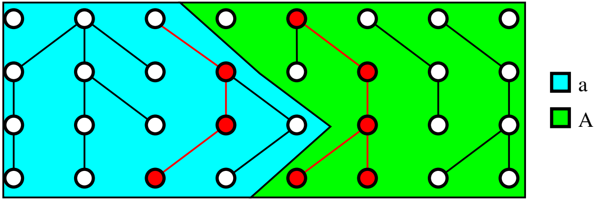
Combining forward processes with the Coalescent

Start with basic coalescent:



Coalescence rate $T_C(k) \simeq \frac{1}{2N} k(\tau)(k(\tau)-1)$

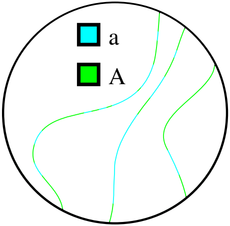
Look closer:



Coalescence rate of a $T_{C,a}(\vec{k}) \simeq \frac{1}{2n_a(\tau)} k_a(k_a-1)$

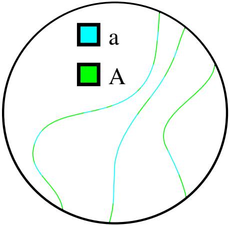
Combining forward processes with the Coalescent

Mutation occurs as a Poisson process along lineages:



Combining forward processes with the Coalescent

Mutation occurs as a Poisson process along lineages:



Retrodictive master equation for lineage mutation:

$$\begin{aligned}\frac{d}{d\tau}R(\vec{k},\tau|\vec{k}_s) &= \mu_{aA}k_a\left[R(\vec{k}_{-a+A},\tau|\vec{k}_s)-R(\vec{k},\tau|\vec{k}_s)\right]\\ &+\mu_{Aa}k_A\left[R(\vec{k}_{-A+a},\tau|\vec{k}_s)-R(\vec{k},\tau|\vec{k}_s)\right]\\ &-\dot{\Omega}(\tau)R(\vec{k},\tau|\vec{k}_s)\end{aligned}$$

Combining forward processes with the Coalescent

RME for combined coalescent/mutation process

$$\begin{aligned}
\frac{d}{d\tau} R(\vec{k}, \tau | \vec{k}_s) \quad = \quad & T_{C,a}(\vec{k}_{+a}) R(\vec{k}_{+a}, \tau | \vec{k}_s) - T_{C,a}(\vec{k}) R(\vec{k}, \tau | \vec{k}_s) \\
& + T_{C,A}(\vec{k}_{+A}) R(\vec{k}_{+A}, \tau | \vec{k}_s) - T_{C,A}(\vec{k}) R(\vec{k}, \tau | \vec{k}_s) \\
& + \mu_{aA} k_a \left[R(\vec{k}_{-a+A}, \tau | \vec{k}_s) - R(\vec{k}, \tau | \vec{k}_s) \right] \\
& + \mu_{Aa} k_A \left[R(\vec{k}_{-A+a}, \tau | \vec{k}_s) - R(\vec{k}, \tau | \vec{k}_s) \right] \\
& - \dot{\Omega}(\tau) R(\vec{k}, \tau | \vec{k}_s)
\end{aligned}$$

Combining forward processes with the Coalescent

Bad News

There are two *unholy* stumbling blocks stopping us from using this approach to directly generate samples from the posterior distribution of trees:

Combining forward processes with the Coalescent

Bad News

There are two *unholy* stumbling blocks stopping us from using this approach to directly generate samples from the posterior distribution of trees:

1. the presence of the background population sizes of each allele in the RME, and

Combining forward processes with the Coalescent

Bad News

There are two *unholy* stumbling blocks stopping us from using this approach to directly generate samples from the posterior distribution of trees:

1. the presence of the background population sizes of each allele in the RME, and
2. the difficulty of simulating coalescence events when sequence space is large.

Conclusions

Conclusions

- Retrodictive Master Equations provide a systematic means of performing retrodictive inference on continuous-time birth/death processes.

Conclusions

- Retrodictive Master Equations provide a systematic means of performing retrodictive inference on continuous-time birth/death processes.
- Have found that solutions to RMEs may be obtained using a ‘retrodictive’ stochastic simulation algorithm.

Conclusions

- Retrodictive Master Equations provide a systematic means of performing retrodictive inference on continuous-time birth/death processes.
- Have found that solutions to RMEs may be obtained using a 'retrodictive' stochastic simulation algorithm.
- Although the RSSA seems like a promising new tool for phylogenetics, no 'killer app.' has yet been found.

Conclusions

- Retrodictive Master Equations provide a systematic means of performing retrodictive inference on continuous-time birth/death processes.
- Have found that solutions to RMEs may be obtained using a 'retrodictive' stochastic simulation algorithm.
- Although the RSSA seems like a promising new tool for phylogenetics, no 'killer app.' has yet been found.

Thank-you!