Viral Phylogeography

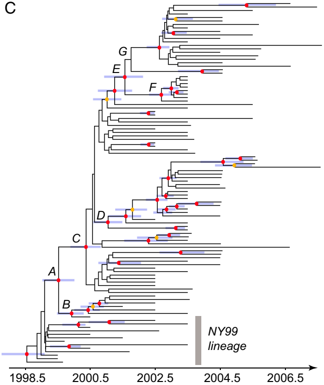
Paul Bastide[1], Gilles Didier[2], Stéphane Guindon[3]

[1] MAP5, Université Paris Cité, CNRS
[2] IMAG, Université de Montpellier, CNRS
[3] LIRMM, Université de Montpellier, CNRS

Rencontre de la chaire MMB, 29 janvier 2025

West Nile Virus (WNV)

(Pybus et al., 2012)



Bayesian Phylogenetics

Goal:

$$p\left(\boldsymbol{\theta}, \mathcal{T}, \psi \mid \mathbf{Y}, \mathbf{S}\right) \propto p\left(\mathbf{Y}, \mathbf{S} \mid \boldsymbol{\theta}, \mathcal{T}, \psi\right) p\left(\boldsymbol{\theta}, \mathcal{T}, \psi\right)$$

Bayesian Phylogenetics

Goal:

$$
\begin{aligned}
p(\boldsymbol{\theta}, \mathcal{T}, \psi \mid \mathbf{Y}, \mathbf{S}) &\propto p(\mathbf{Y}, \mathbf{S} \mid \boldsymbol{\theta}, \mathcal{T}, \psi)\, p(\boldsymbol{\theta}, \mathcal{T}, \psi) \\
&\propto p(\mathbf{Y} \mid \boldsymbol{\theta}, \mathcal{T})\, p(\mathbf{S} \mid \mathcal{T}, \psi)\, p(\boldsymbol{\theta}, \mathcal{T}, \psi)
\end{aligned}
$$

Assumption: $\mathbf{Y}$ and $\mathbf{S}$ independent conditionally on $\mathcal{T}$.

Bayesian Phylogenetics

Goal:

$$
\begin{aligned}
p(\boldsymbol{\theta}, \mathcal{T}, \psi \mid \mathbf{Y}, \mathbf{S}) &\propto p(\mathbf{Y}, \mathbf{S} \mid \boldsymbol{\theta}, \mathcal{T}, \psi)\, p(\boldsymbol{\theta}, \mathcal{T}, \psi) \\
&\propto p(\mathbf{Y} \mid \boldsymbol{\theta}, \mathcal{T})\, p(\mathbf{S} \mid \mathcal{T}, \psi)\, p(\boldsymbol{\theta}, \mathcal{T}, \psi) \\
&\propto p(\mathbf{Y} \mid \boldsymbol{\theta}, \mathcal{T})\, p(\boldsymbol{\theta})\, p(\mathbf{S} \mid \mathcal{T}, \psi)\, p(\mathcal{T}, \psi)
\end{aligned}
$$

Assumption: $\mathbf{Y}$ and $\mathbf{S}$ independent conditionally on $\mathcal{T}$.

This talk: Conditional on $\mathcal{T}$.

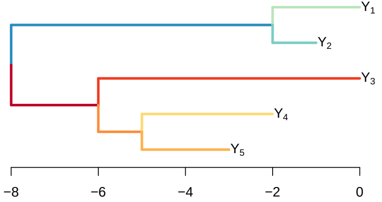
Outline

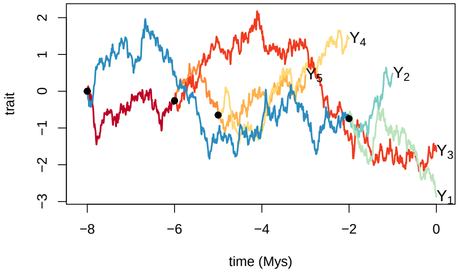
1. From the Brownian Motion to the Cauchy Process
    - Brownian Motion
    - Relaxed Brownian Motion
    - Cauchy Process
2. Cauchy Process on a Tree
    - CP on a Tree
    - Likelihood Computation
    - Ancestral State Reconstruction
3. Integrated Processes
    - Velocity Statistic
    - Integrated Brownian Motion
    - Belief Propagation

Outline

1. From the Brownian Motion to the Cauchy Process
    - Brownian Motion
    - Relaxed Brownian Motion
    - Cauchy Process
2. Cauchy Process on a Tree
3. Integrated Processes

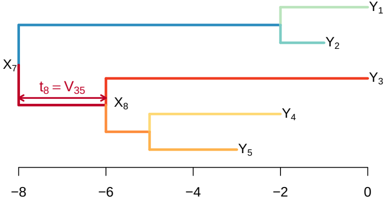
Brownian Motion on a Tree (Felsenstein, 1985)

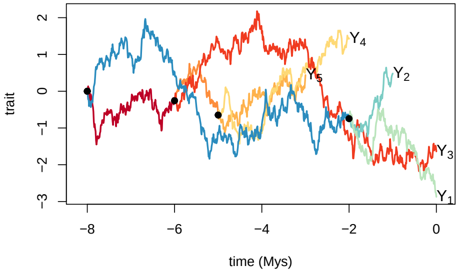




- The trait evolves like a BM in time
- Speciation → two independent processes
- Only tip values are measured

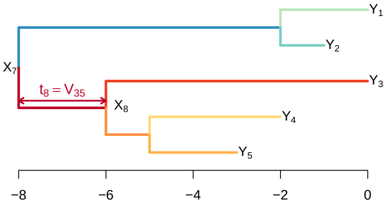
Brownian Motion on a Tree (Felsenstein, 1985)

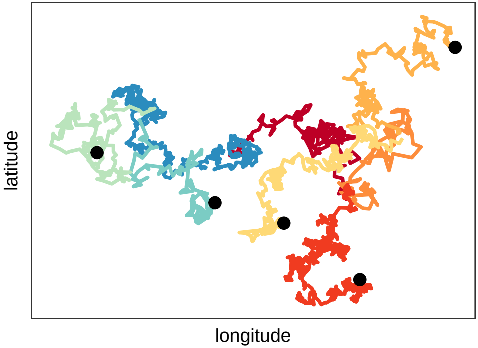




- SDE: $\mathrm{d}\,X_t = \sigma\,\mathrm{d}\,B_t$
- Heredity: $X_8 | X_7 \sim \mathcal{N}(X_7, \sigma^2 t_8)$
- Covariances: $\mathbb{C}\mathrm{ov}(Y_i, Y_j) = \sigma^2 V_{ij}$
- Distribution: $\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma^2 \mathbf{V})$

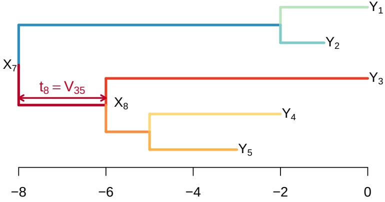
Bi-variate Brownian Motion on a Tree

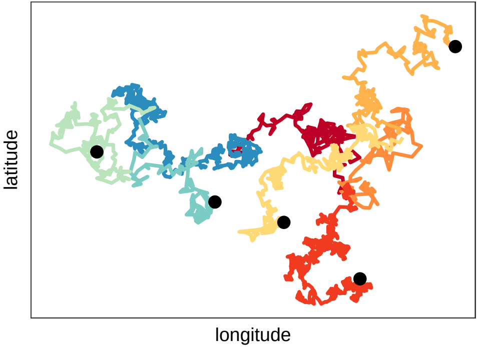




- SDE: $\mathrm{d}\,\mathbf{X}_t = \mathbf{\Sigma}^{1/2}\,\mathrm{d}\,\mathbf{B}_t$
- Heredity: $\mathbf{X}_8|\mathbf{X}_7 \sim \mathcal{N}(\mathbf{X}_7, t_8\mathbf{\Sigma})$
- Distribution: $\mathbf{Y} \sim \mathcal{MN}(\mathbf{1}_n\boldsymbol{\mu}^T, \mathbf{V}, \mathbf{\Sigma})$

Bi-variate Brownian Motion on a Tree

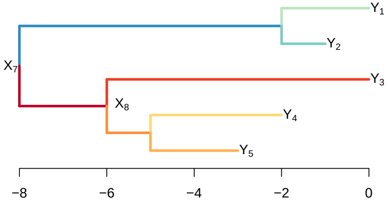




- SDE: $\mathrm{d}\,\mathbf{X}_t = \mathbf{\Sigma}^{1/2}\,\mathrm{d}\,\mathbf{B}_t$
- Heredity: $\mathbf{X}_8|\mathbf{X}_7 \sim \mathcal{N}(\mathbf{X}_7, t_8\mathbf{\Sigma})$
- Distribution: $\mathbf{Y} \sim \mathcal{MN}(\mathbf{1}_n\boldsymbol{\mu}^T, \mathbf{V}, \mathbf{\Sigma})$
- Note: $\mathbf{Y}_3|\mathbf{X}_7 \sim \mathcal{N}(X_7, (t_8 + t_3)\mathbf{\Sigma})$

Relaxed Random Walk

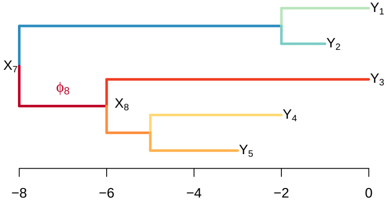
(Lemey et al., 2010)



- Brownian Motion: $\mathbf{X}_8 | \mathbf{X}_7 \sim \mathcal{N}(\mathbf{X}_7, t_8 \mathbf{\Sigma})$

Relaxed Random Walk

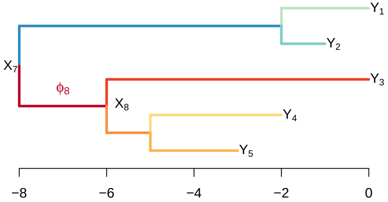
(Lemey et al., 2010)



- Brownian Motion: $\mathbf{X}_8 | \mathbf{X}_7 \sim \mathcal{N}(\mathbf{X}_7, t_8 \boldsymbol{\Sigma})$
- Relaxed Brownian Motion: $\mathbf{X}_8 | \mathbf{X}_7, \phi_8 \sim \mathcal{N}(\mathbf{X}_7, \phi_8 \times t_8 \boldsymbol{\Sigma})$

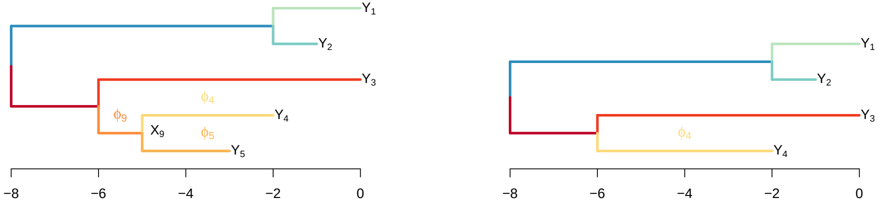
Relaxed Random Walk

(Lemey et al., 2010)



- Brownian Motion: $\mathbf{X}_8 | \mathbf{X}_7 \sim \mathcal{N}(\mathbf{X}_7, t_8 \mathbf{\Sigma})$
- Relaxed Brownian Motion: $\mathbf{X}_8 | \mathbf{X}_7, \phi_8 \sim \mathcal{N}(\mathbf{X}_7, \phi_8 \times t_8 \mathbf{\Sigma})$
- Regularisation: $\phi_j \sim \mathcal{L}(\boldsymbol{\theta})$ iid

RRW: Not Sampling Consistent



- Relaxed Brownian Motion: $\mathbf{X}_i | \mathbf{X}_{\mathrm{pa}(i)}, \phi_i \sim \mathcal{N}(\mathbf{X}_{\mathrm{pa}(i)}, \phi_i \times t_i \boldsymbol{\Sigma})$
- Regularisation: $\phi_j \sim \mathcal{L}(\boldsymbol{\theta})$ iid
- Unsampled tip: Changes the whole distribution.

Inverse-Gamma Normal Mixture

$$\phi_j \mid \nu \sim \text{Inv-Gamma}(\nu/2, \nu/2)$$
$$\mathbf{X}_j \mid \mathbf{X}_{\text{pa}(j)}, \phi_j, \mathbf{\Sigma} \sim \mathcal{N}(\mathbf{X}_{\text{pa}(j)}, \phi_j \times t_j \mathbf{\Sigma})$$

Inverse-Gamma Normal Mixture

$$\phi_j \mid \nu \sim \text{Inv-Gamma}(\nu/2, \nu/2)$$
$$\mathbf{X}_j \mid \mathbf{X}_{\text{pa}(j)}, \phi_j, \mathbf{\Sigma} \sim \mathcal{N}(\mathbf{X}_{\text{pa}(j)}, \phi_j \times t_j \mathbf{\Sigma})$$

gives:

$$\mathbf{X}_j \mid \mathbf{X}_{\text{pa}(j)}, \mathbf{\Sigma}, \nu \sim \mathcal{T}_\nu(\mathbf{X}_{\text{pa}(j)}, t_j \mathbf{\Sigma})$$

Inverse-Gamma Normal Mixture

$$\phi_j \mid \nu \sim \text{Inv-Gamma}(\nu/2, \nu/2)$$
$$\mathbf{X}_j \mid \mathbf{X}_{\text{pa}(j)}, \phi_j, \mathbf{\Sigma} \sim \mathcal{N}(\mathbf{X}_{\text{pa}(j)}, \phi_j \times t_j \mathbf{\Sigma})$$

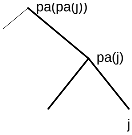
gives:

$$\mathbf{X}_j \mid \mathbf{X}_{\text{pa}(j)}, \mathbf{\Sigma}, \nu \sim \mathcal{T}_\nu(\mathbf{X}_{\text{pa}(j)}, t_j \mathbf{\Sigma})$$

Student: not a stable distribution.

$$\mathbf{X}_j \mid \mathbf{X}_{\text{pa}(\text{pa}(j))}, \mathbf{\Sigma}, \nu \nsim \mathcal{T}_\nu(\mathbf{X}_{\text{pa}(\text{pa}(j))}, (t_j + t_{pa(j)}) \mathbf{\Sigma})$$

$\rightarrow$ adding a node changes the whole distribution.



Stable Distributions

Stable distribution: X is stable if, for X_1 and X_2 iid copies of X:

$$aX_1 + bX_2 \underset{\text{dist}}{\equiv} cX + d \qquad (a, b, c > 0)$$

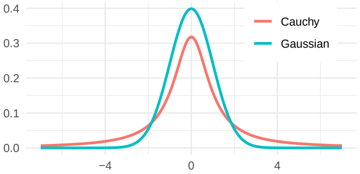
Characteristic function: Symmetric α stable distributions :

$$\mathbb{E}\left[e^{iuX}\right] = \phi(u; \alpha, \gamma, \mu) = \exp(iu\mu - |\gamma u|^{\alpha})$$

Density: Only tractable in two special cases:

- $\alpha = 2$: Gaussian distribution.
- $\alpha = 1$: Cauchy distribution.

Note: See also the Lévy distribution for an *asymmetric* stable distribution.



Cauchy Process (CP)

Cauchy Propagation:

$$X_j \mid X_{\mathrm{pa}(j)}, \sigma \sim \mathcal{C}(X_{\mathrm{pa}(j)}, \sigma t_j)$$

Stable Distribution:

$$X_j \mid X_{\mathrm{pa}(\mathrm{pa}(j))}, \sigma \sim \mathcal{C}(X_{\mathrm{pa}(\mathrm{pa}(j))}, \sigma(t_j + t_{\mathrm{pa}(j)}))$$

Density:

$$p\left(X_j \mid X_{\mathrm{pa}(j)}, \sigma\right) = \frac{1}{\pi \sigma t_j} \frac{1}{1 + \left(\frac{X_j - X_{\mathrm{pa}(j)}}{\sigma t_j}\right)^2}$$

- $X_{\mathrm{pa}(j)}$ is the **location parameter**.
- σt_j is the **scale parameter**.

Inverse-Gamma Normal Mixture

$$\phi_j \mid \nu \sim \text{Inv-Gamma}(1/2, 1/2)$$
$$X_j \mid X_{\text{pa}(j)}, \phi_j, \sigma^2 \sim \mathcal{N}(X_{\text{pa}(j)}, \phi_j \times \sigma^2 t_j)$$

gives:

$$X_j \mid X_{\text{pa}(j)}, \sigma^2 \sim \mathcal{T}_1(X_{\text{pa}(j)}, \sigma^2 t_j)$$

Inverse-Gamma Normal Mixture

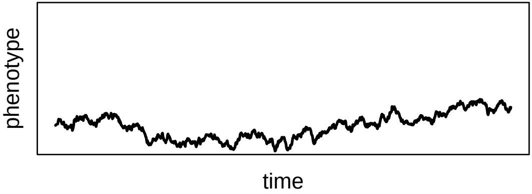
$$\begin{aligned} \phi_j \mid \nu &\sim \text{Inv-Gamma}(1/2, 1/2) \\ X_j \mid X_{\text{pa}(j)}, \phi_j, \sigma^2 &\sim \mathcal{N}(X_{\text{pa}(j)}, \phi_j \times \sigma^2 t_j) \end{aligned}$$

gives:

$$\begin{aligned} X_j \mid X_{\text{pa}(j)}, \sigma^2 &\sim \mathcal{T}_1(X_{\text{pa}(j)}, \sigma^2 t_j) \\ &\sim \mathcal{C}(X_{\text{pa}(j)}, \sigma \sqrt{t_j}) \end{aligned}$$

"Cauchy-RRW" $\rightarrow$ Still not stable !

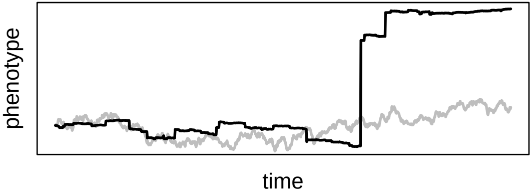
Pure Jump Lévy Process



Brownian Motion:

$$\mathbb{E}\left[e^{iuX(t)}\right] = e^{-t\frac{\sigma^2}{2}u^2}$$

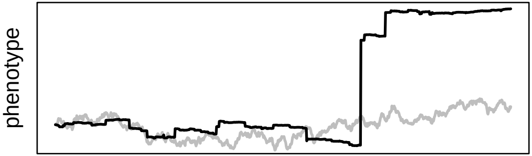
Pure Jump Lévy Process



Cauchy Process:

$$\mathbb{E}\left[e^{iuX(t)}\right] = \exp\left(\frac{\sigma t}{\pi}\int_{\mathbb{R}}\left(e^{iux} - 1 - iux\mathbb{I}\{|x| < 1\}\right)\frac{\mathrm{d}\,x}{x^2}\right) = e^{-\sigma t|u|}$$

Pure Jump Lévy Process



Cauchy Process:

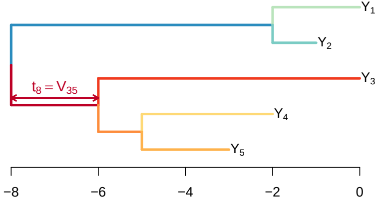
$$X_t^{Cau} = X_t^0 + \sum_{k \geq 1} (X_t^k - \mathbb{E}[X_t^k])$$

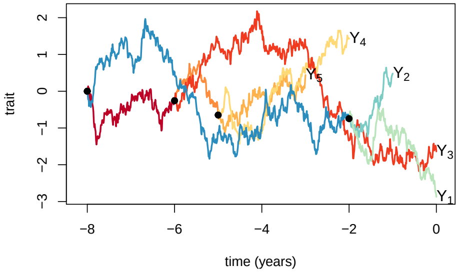
$$X_t^k \sim \text{Compound Poisson} \qquad \text{rate: } \frac{2\sigma}{\pi} \qquad \text{dist: } \frac{\mathrm{d}\,x}{2x^2} \text{ on } I_k$$

$$I_0 =]-\infty; -1] \cup [1; +\infty[, \; I_k = \left]-\frac{1}{k}; -\frac{1}{k+1}\right] \cup \left[\frac{1}{k+1}; \frac{1}{k}\right[\; k \geq 1$$

Outline

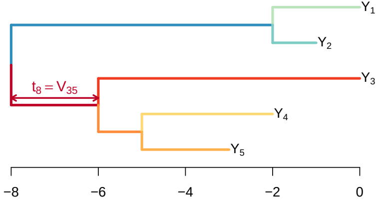
CP on a Tree

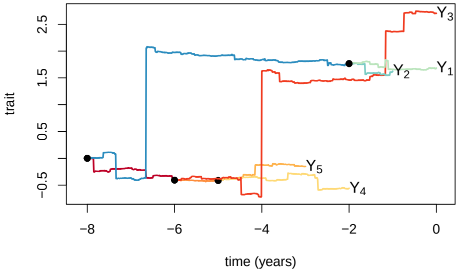




- Heredity: $X_8 | X_7 \sim \mathcal{N}(X_7, \sigma^2 t_8)$
- Covariances: $\mathbb{C}\text{ov}(Y_i, Y_j) = \sigma^2 V_{ij}$
- Marginal: $Y_i \sim \mathcal{N}(\mu, \sigma^2 V_{ii})$
- Distribution: $\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma^2 \mathbf{V})$

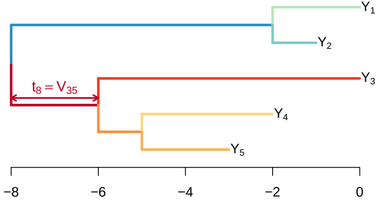
CP on a Tree

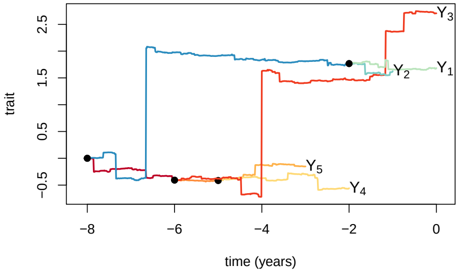




- Heredity: $X_8 | X_7 \sim \mathcal{C}(X_7, \sigma t_8)$
- Covariances: do not exist.
- Marginal: $Y_i \sim \mathcal{C}(\mu, \sigma \tau_i)$
- Distribution: $\mathbf{Y} \sim ?$

CP on a Tree





- **Heredity:** $X_8|X_7 \sim \mathcal{C}(X_7, \sigma t_8)$
- **Covariances:** do not exist.
- **Marginal:** $Y_i \sim \mathcal{C}(\mu, \sigma\tau_i)$
- **Distribution:** $\mathbf{Y} \sim \mathcal{MC}(\mu\mathbf{1}, \gamma_{phy}(\cdot))$

Multivariate Cauchy

Likelihood Computation

Likelihood: Characteristic Function

$$p(\mathbf{Y} \mid \mu, \sigma) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\mathbf{u}^T\mathbf{Y}} \phi_{\mathbf{Y}|X_{\text{root}}}(\mathbf{u}; \sigma) \, \mathrm{d}\,\mathbf{u}$$

Exact Algorithm: Algo

Can compute the integral explicitly, with one traversal of the tree.

Complexity:

Quadratic in the number of tips.

Stability:

Sums of large positive and negative numbers: numerical issues.

Ancestral State Reconstruction

Density for an ancestral state:

$$p(X_j = v \mid \mathbf{Y}, X_r = \mu, \sigma) = \frac{p(\mathbf{Y}^j \mid X_j = v, \sigma) p(\mathbf{Y}^{-j}, X_j = v \mid X_r = \mu, \sigma)}{p(\mathbf{Y} \mid X_r = \mu, \sigma)}$$

Can be computed for a grid of values v.
(Linear in the number of grid values.)

Ancestral State Reconstruction

Density for an ancestral state:

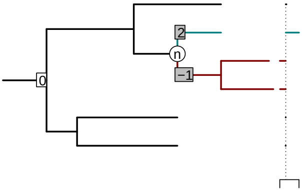
$$p(X_j = v \mid \mathbf{Y}, X_r = \mu, \sigma) = \frac{p(\mathbf{Y}^j \mid X_j = v, \sigma) p(\mathbf{Y}^{-j}, X_j = v \mid X_r = \mu, \sigma)}{p(\mathbf{Y} \mid X_r = \mu, \sigma)}$$

Can be computed for a grid of values v.
(Linear in the number of grid values.)

Can be multimodal !

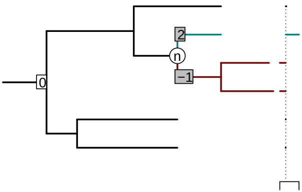
Identifiability Issues

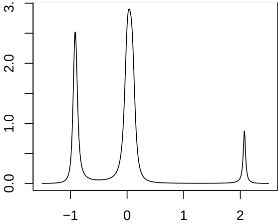
Cauchy reconstruction:



Identifiability Issues

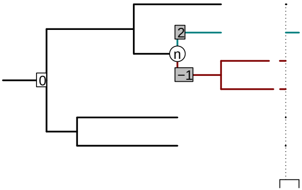
Cauchy reconstruction:

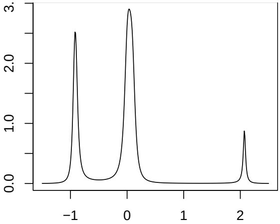




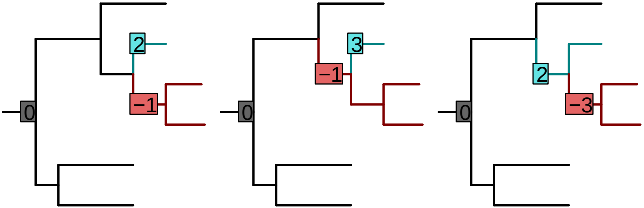
Identifiability Issues

Cauchy reconstruction:

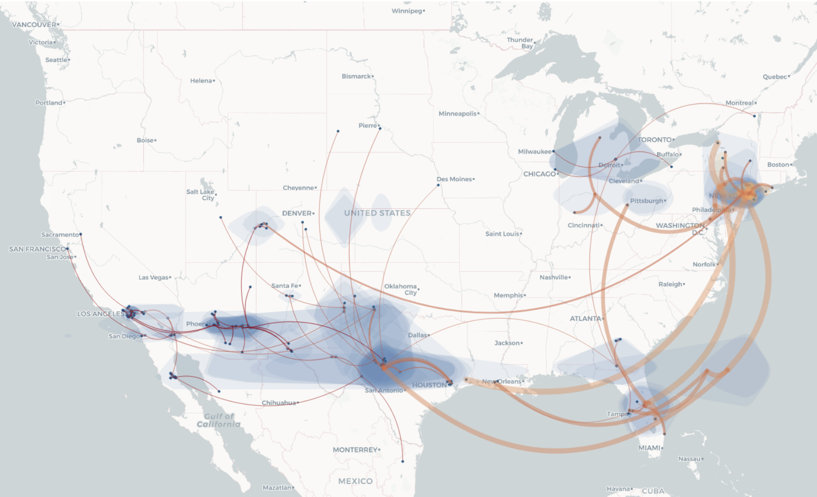




Shifted BM: Some shift configuration are not identifiable.

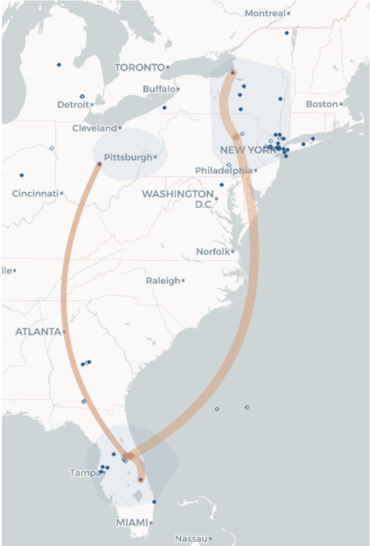


CP:


VANCOUVER
Victoria
Seattle
Portland
Helena
Bismarck
Winnipeg
Thunder Bay
Quebec
Montreal
Minneapolis
Pierre
Boise
Milwaukee
TORONTO
Buffalo
Detroit
Boston
CHICAGO
Cleveland
Des Moines
Cheyenne
Salt Lake City
Pittsburgh
DENVER
UNITED STATES
Philadelphia
Cincinnati
WASHINGTON D.C.
Saint Louis
Sacramento
SAN FRANCISCO
San Jose
Norfolk
Las Vegas
Nashville
Raleigh
Santa Fe
Oklahoma City
Memphis
LOS ANGELES
ATLANTA
Phoenix
Dallas
San Diego
Jackson
HOUSTON
New Orleans
San Antonio
Chihuahua
Tampa
Gulf of California
MIAMI
Nassau
MONTERREY
MEXICO
Mazatlán
Havana
CUBA

WNV using Evolaps (Pybus et al., 2012; Chevenet et al., 2024)

CP:

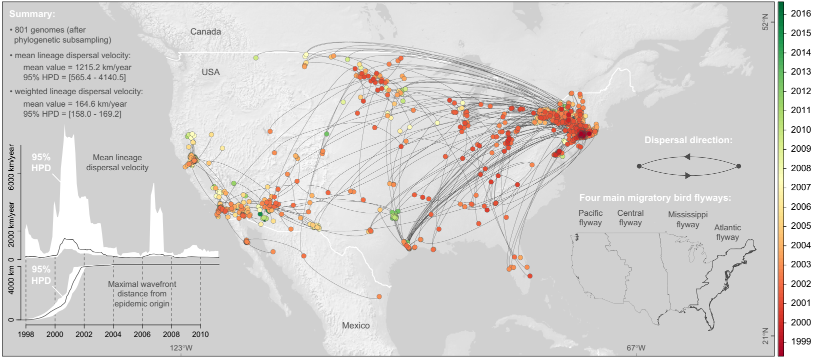


Cauchy RRW:

Simulations

Outline

How fast is the virus going ? (Dellicour et al., 2020)



Velocity Statistic:

$$WLDV = \frac{\sum_{i \in \text{branches}} d_i}{\sum_{i \in \text{branches}} t_i}$$

Velocity Statistic

MCMC: Sample $(\boldsymbol{\theta}_k, \mathcal{T}_k, \boldsymbol{\psi}_k)$ from

$$p\left(\boldsymbol{\theta}, \mathcal{T}, \boldsymbol{\psi} \mid \mathbf{Y}, \mathbf{S}\right)$$

Ancestral reconstruction: Sample ancestral positions $\mathbf{X}_k$ from

$$p\left(\mathbf{X} \mid \mathbf{Y}, \boldsymbol{\theta}_k, \mathcal{T}_k\right)$$

Displacement: for each branch i at iteration k

$$d_k^i = \|X_k^{\mathrm{pa}(i)} - X_k^i\|_2 = \text{distance covered on branch } i$$
$$t_k^i = \text{length of branch } i \text{ in tree } \mathcal{T}_k$$

Velocity Statistic:

$$WLDV_k = \frac{\sum_{i=1}^N d_k^i}{\sum_{i=1}^N t_k^i}$$

Problem with the Velocity Statistic (Dellicour et al., 2024)

Models: Brownian Motion, Relaxed Random Walk, Cauchy

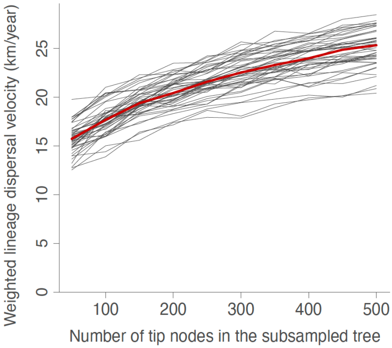
Problem with the Velocity Statistic (Dellicour et al., 2024)

Models: Brownian Motion, Relaxed Random Walk, Cauchy

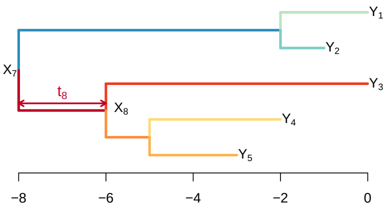
Velocity is not defined ! processes have infinite variation.

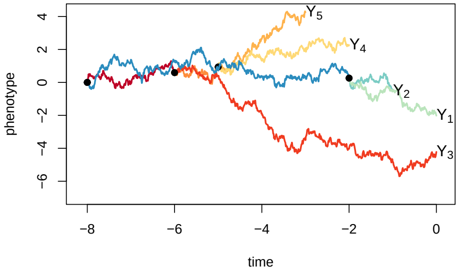
Problem with the Velocity Statistic (Dellicour et al., 2024)

Models: Brownian Motion, Relaxed Random Walk, Cauchy
Velocity is not defined ! processes have infinite variation.
Sampling inconsistent: the more densely sampled, the faster the process appears



Brownian Motion

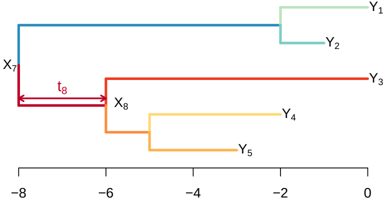


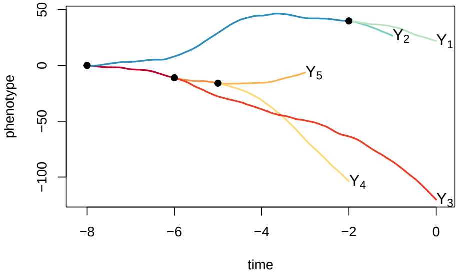


Brownian Motion

$$X(t) = x_0 + \int_0^t \sigma dB_t \qquad \sim \mathcal{N}(x_0, \sigma^2 t)$$

Integrated Brownian Motion



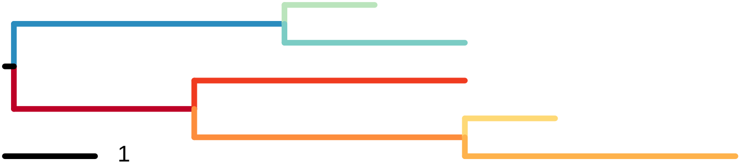


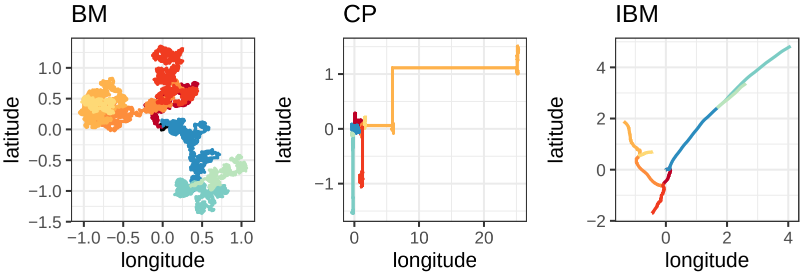
Integrated Brownian Motion

$$V(t) = v_0 + \int_0^t \sigma dB_t \qquad \sim \mathcal{N}(v_0, \sigma^2 t)$$

$$X(t) = x_0 + \int_0^t V(s) ds \qquad \sim \mathcal{N}(x_0 + v_0 t, \frac{\sigma^2}{3} t^3)$$

Integrated Brownian Motion





Integrated Brownian Motion is Gaussian

Integrated Brownian Motion

$$\mathbf{V}(t) = \mathbf{V}_\rho + \int_0^t \sigma d\mathbf{B}_t \qquad \mathbf{X}(t) = \mathbf{X}_\rho + \int_0^t \mathbf{V}(s) ds$$

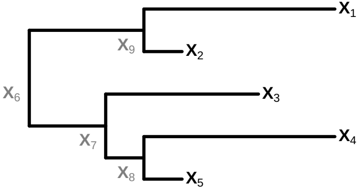
Velocity (BM):

$$\mathbb{E}[\mathbf{V}_i] = \mathbf{V}_\rho$$
$$\mathbb{V}\text{ar}[\mathbf{V}_i; \mathbf{V}_j] = \mathbf{\Sigma}\tau_{ij}$$

Position (IBM):

$$\mathbb{E}[\mathbf{X}_i] = \mathbf{V}_\rho \tau_i + \mathbf{X}_\rho$$
$$\mathbb{V}\text{ar}[\mathbf{X}_i; \mathbf{X}_j] = \mathbf{\Sigma}\tau_{ij} \left[\tau_i\tau_j + \tau_{ij}\left(\frac{\tau_{ij}}{3} - \frac{\tau_i + \tau_j}{2}\right)\right].$$

Efficient Computation



$$\mathbf{X}_r \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Gamma})$$

$$\mathbf{X}_j \mid \mathbf{X}_{\text{pa}(j)} \sim \mathcal{N}\left(\mathbf{q}_j \mathbf{X}_{\text{pa}(j)} + \mathbf{r}_j,\ \boldsymbol{\Sigma}_j\right)$$

Pruning: Likelihood $p_\theta(\mathbf{X}_{obs})$ in $O(n)$.

Belief Propagation: Ancestral density $p_\theta(\mathbf{X}_{anc}|\mathbf{X}_{obs})$ in $O(n)$.

$\rightarrow$ with missing data.

Efficient Computation

BM:

$$\mathbf{X}_j \mid \mathbf{X}_{\text{pa}(j)} \sim \mathcal{N}\left(\mathbf{X}_{\text{pa}(j)},\ t_j \mathbf{R}\right)$$

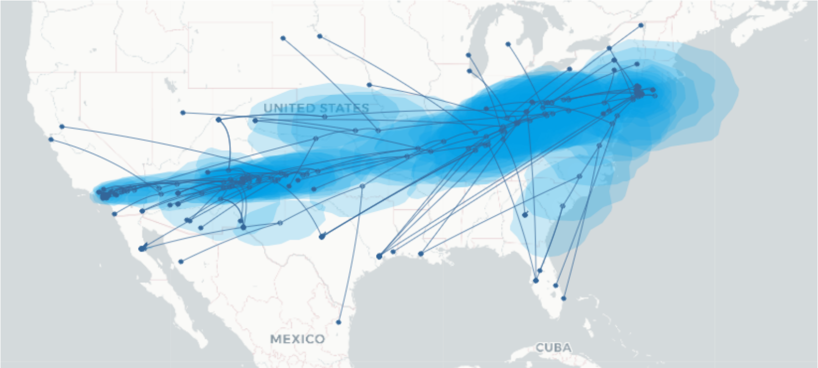
IBM:

$$\begin{pmatrix}\mathbf{V}_j \\ \mathbf{X}_j\end{pmatrix} \Bigm| \begin{pmatrix}\mathbf{V}_{\text{pa}(j)} \\ \mathbf{X}_{\text{pa}(j)}\end{pmatrix} \sim$$

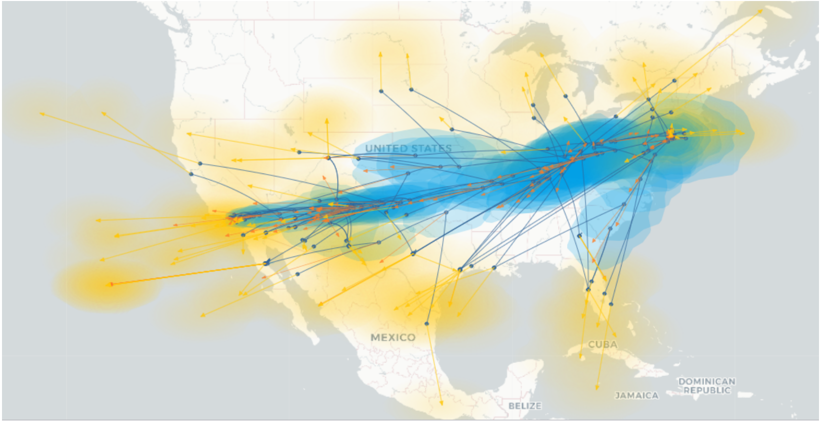
$$\mathcal{N}\left(\left[\begin{pmatrix}1 & 0 \\ t_i & 1\end{pmatrix} \otimes \mathbf{I}_p\right] \begin{pmatrix}\mathbf{V}_{\text{pa}(j)} \\ \mathbf{X}_{\text{pa}(j)}\end{pmatrix},\ \begin{pmatrix}t_i & t_i^2/2 \\ t_i^2/2 & t_i^3/3\end{pmatrix} \otimes \mathbf{\Sigma}\right)$$

→ joint velocity / position vector is linear Gaussian

Integrated Brownian Motion



Integrated Brownian Motion - Predictions ?



Perspectives

Three processes:

- Strict Brownian Motion
- Relaxed Random Walk - Cauchy Process
- Integrated BM IOU

Efficient Likelihood Computation:

- Linear Gaussian process
- Sequence / Trait independence
- No geography / temporal variables

Perspectives:

- Multivariate Cauchy +
- Include spacial co-variables
- Prediction ?

References

P. Bastide, L.S.T. Ho, G. Baele, P. Lemey, M.A. Suchard (2021), Efficient Bayesian Inference of General Gaussian Models on Large Phylogenetic Trees. *Annals of Applied Statistics*, 15(2): 971-997. Implemented in `BEAST`.

P. Bastide, G. Didier (2023), The Cauchy Process on Phylogenies: a Tractable Model for Pulsed Evolution. *Systematic Biology*, 72(6), 1296–1315. `R` package `cauphy` on GitHub.

S. Dellicour, P. Bastide, P. Rocu, D. Fargette, O.J. Hardy, M.A. Suchard, S. Guindon, P. Lemey (2024), How fast are viruses spreading in the wild? *PLOS Biology* (to appear).

P. Bastide, P. Rocu, J. Wirtz, G.W. Hassler, F. Chevenet, D. Fargette, M.A. Suchard, S. Dellicour, P. Lemey, S. Guindon (2024), Modeling the velocity of evolving lineages and predicting dispersal patterns. *P.N.A.S.*, 121 (47) e2411582121. Implemented in `BEAST`.

Bibliography I

Chevenet, Fargette, Bastide, et al. 2024. *Virus Evolution*. 10.

Dellicour, Bastide, Rocu, et al. 2024. *PLOS Biology*. 22:e3002914.

Dellicour, Lequime, Vrancken, et al. 2020. *Nature Communications*. 11.

Duchen, Leuenberger, Szilágyi, et al. 2017. *Systematic Biology*. 66:950–963.

Elliot, Mooers. 2014. *BMC Evolutionary Biology*. 14:226.

Felsenstein. 1985. *The American Naturalist*. 125:1–15.

Ferguson. 1962. *The Annals of Mathematical Statistics*. 33:1256–1266.

Fisher, Ji, Lemey, et al. 2021. *Systematic Biology*. 70:258–267.

Jaffe, Slater, Alfaro. 2011. *Biology Letters*. 7:558–561.

Kidmose. 2001. In: IEEE-EURASIP 2001 Workshop on Nonlinear Signal and Image Processing.

Landis, Schraiber. 2017. *Proceedings of the National Academy of Sciences*. 114:13224–13229.

Landis, Schraiber, Liang. 2013. *Systematic Biology*. 62:193–204.

Lemey, Rambaut, Welch, et al. 2010. *Molecular Biology and Evolution*. 27:1877–1885.

Leventhal, Bonhoeffer. 2016. *Trends in Microbiology*. 24:687–698.

Bibliography II

Mahler, Ingram, Revell, et al. 2013. *Science*. 341:292–295.

Nolan. 2005. In: Deutsche Bundesbank's 2005 annual fall conference.

Pagel. 1999. *Nature*. 401:877–884.

Pybus, Suchard, Lemey, et al. 2012. *Proceedings of the National Academy of Sciences*. 109:15066–15071.

Photo Credits

Thank you for listening

Gilles Didier
IMAG, Université de Montpellier, CNRS

Stéphane Guindon, Pauline Rocu, Johannes Wirtz
LIRMM, Université de Montpellier, CNRS

Denis Fargette
PHIM, Université de Montpellier, IRD

François Chevenet, Thomas Vitré
MIVEGEC, Université de Montpellier, IRD

Philippe Lemey
Evolutionary and Computational Virology, KU Leuven

Simon Dellicour
Spatial Epidemiology Lab, ULB

Gabriel Hassler, Marc Suchard
Biomathematics, Biostatistics and Human Genetics
University of California, Los Angeles

Appendices

Multivariate Cauchy (Ferguson, 1962)

Definition:

$$\mathbf{X} \sim \mathcal{MC}_p \iff \mathbf{a}^T\mathbf{X} \sim \mathcal{C} \quad , \; \forall \mathbf{a} \in \mathbb{R}^p.$$

Multivariate Cauchy (Ferguson, 1962)

Definition:

$$\mathbf{X} \sim \mathcal{MC}_p \iff \mathbf{a}^T\mathbf{X} \sim \mathcal{C} \quad, \ \forall \mathbf{a} \in \mathbb{R}^p.$$

Characterization:

$$\mathbf{X} \sim \mathcal{MC}_p \iff \phi_{\mathbf{X}}(\mathbf{u}) = \mathbb{E}\left[e^{\mathbf{u}^T\mathbf{X}}\right] = e^{i\mu(\mathbf{u}) - \gamma(\mathbf{u})}$$

with:

$$\begin{cases} \mu(a\mathbf{u}) = a\mu(\mathbf{u}) \\ \gamma(a\mathbf{u}) = |a|\,\gamma(\mathbf{u}) \end{cases} \qquad \forall a \in \mathbb{R}, \mathbf{u} \in \mathbb{R}^p.$$

Multivariate Cauchy (Ferguson, 1962)

Definition:

$$\mathbf{X} \sim \mathcal{MC}_p \iff \mathbf{a}^T\mathbf{X} \sim \mathcal{C} \quad , \ \forall \mathbf{a} \in \mathbb{R}^p.$$

Characterization:

$$\mathbf{X} \sim \mathcal{MC}_p \iff \phi_{\mathbf{X}}(\mathbf{u}) = \mathbb{E}\left[e^{\mathbf{u}^T\mathbf{X}}\right] = e^{i\mu(\mathbf{u}) - \gamma(\mathbf{u})}$$

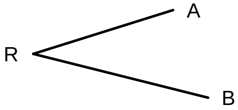
with:

$$\begin{cases} \mu(a\mathbf{u}) = a\mu(\mathbf{u}) \\ \gamma(a\mathbf{u}) = |a|\,\gamma(\mathbf{u}) \end{cases} \qquad \forall a \in \mathbb{R}, \mathbf{u} \in \mathbb{R}^p.$$

Student with $\nu = 1$:

$$\mathbf{X} \sim \mathcal{MT}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}; \nu = 1) \iff \phi_{\mathbf{X}}(\mathbf{u}) = e^{\mathbf{u}^T\mu - \sqrt{\mathbf{u}^T\boldsymbol{\Sigma}\mathbf{u}}}$$

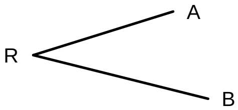
Two tips tree



$$\phi_{Y_A|R}(u_A;\sigma) = \exp\left(i\mu u_A - \sigma t_A |u_A|\right)$$
$$\phi_{Y_B|R}(u_B;\sigma) = \exp\left(i\mu u_B - \sigma t_B |u_B|\right)$$

back

Two tips tree



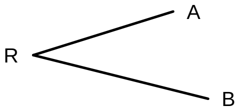
$$\phi_{Y_A|R}(u_A;\sigma) = \exp\left(i\mu u_A - \sigma t_A |u_A|\right)$$
$$\phi_{Y_B|R}(u_B;\sigma) = \exp\left(i\mu u_B - \sigma t_B |u_B|\right)$$

Joint distribution:

$$\begin{aligned}\phi_{Y_A,Y_B|R}(\mathbf{u};\sigma) &= \phi_{Y_A|R}(u_A;\sigma) \times \phi_{Y_B|R}(u_B;\sigma)\\ &= \exp\left(i\mu(u_A + u_B) - \sigma(t_A |u_A| + t_B |u_B|)\right)\end{aligned}$$

back

Two tips tree



$$\phi_{Y_A|R}(u_A;\sigma) = \exp\left(i\mu u_A - \sigma t_A |u_A|\right)$$
$$\phi_{Y_B|R}(u_B;\sigma) = \exp\left(i\mu u_B - \sigma t_B |u_B|\right)$$

Joint distribution:

$$\begin{aligned}\phi_{Y_A,Y_B|R}(\mathbf{u};\sigma) &= \phi_{Y_A|R}(u_A;\sigma) \times \phi_{Y_B|R}(u_B;\sigma)\\ &= \exp\left(i\mu(u_A+u_B) - \sigma(t_A|u_A| + t_B|u_B|)\right)\end{aligned}$$

$\rightarrow$ multivariate Cauchy...

...but not Student:

$$\gamma(\mathbf{u}) = \sigma(t_A|u_A| + t_B|u_B|) \neq \sqrt{\mathbf{u}^T\mathbf{\Sigma}\mathbf{u}}$$

back

Characteristic Function

Branch Characteristic Function:

$$\phi_{X_j | X_{\text{pa}(j)}}(u;\sigma) = \exp\left(iX_{\text{pa}(j)}u - \sigma t_j |u|\right)$$

Tree Characteristic Function: Conditionally on $X_{\text{root}} = \mu$:

$$\phi_{\mathbf{Y}|X_{\text{root}}}(\mathbf{u};\sigma) = \exp\left(i\mu\sum_{k=1}^{n} u_k - \sigma \sum_{j\neq\text{root}} t_j \left|\sum_{k\in\text{desTips}(j)} u_k\right|\right)$$

Characteristic Function

Branch Characteristic Function:

$$\phi_{X_j \middle| X_{\mathrm{pa}(j)}}(u;\sigma) = \exp\left(iX_{\mathrm{pa}(j)}u - \sigma t_j \left|u\right|\right)$$

Tree Characteristic Function: Conditionally on $X_{\mathrm{root}} = \mu$:

$$\phi_{\mathbf{Y}|X_{\mathrm{root}}}(\mathbf{u};\sigma) = \exp\left(i\mu\sum_{k=1}^{n}u_k - \sigma\sum_{j\neq\mathrm{root}}t_j\left|\sum_{k\in\mathrm{desTips}(j)}u_k\right|\right)$$

+

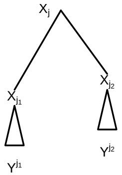
Multivariate Cauchy but **not** a Student with $\nu = 1$.

back

Characteristic Function : Pre-order Tree Traversal

Characteristic Function:

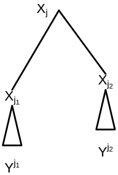
$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \exp\left(iX_j \sum_{k\in \mathrm{desTips}(j)} u_k - \sigma \sum_{k\in \mathrm{des}\, j} t_k \left| \sum_{l\in \mathrm{desTips}(k)} u_l \right| \right)$$



Characteristic Function : Pre-order Tree Traversal

Characteristic Function:

$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \exp\left(iX_j \sum_{k\in \mathrm{desTips}(j)} u_k - \sigma \sum_{k \in \mathrm{des}\, j} t_k \left| \sum_{l \in \mathrm{desTips}(k)} u_l \right| \right)$$



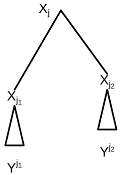
Propagation:

$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right) \times \phi_{\mathbf{Y^{j_2}}|X_j}\left(\mathbf{u}^{j_2};\sigma\right)$$ conditional independence

Characteristic Function : Pre-order Tree Traversal

Characteristic Function:

$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \exp\left(iX_j \sum_{k\in \text{desTips}(j)} u_k - \sigma \sum_{k \in \text{des}\, j} t_k \left| \sum_{l\in\text{desTips}(k)} u_l \right|\right)$$



Propagation:

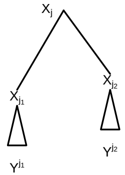
$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right) \times \phi_{\mathbf{Y^{j_2}}|X_j}\left(\mathbf{u}^{j_2};\sigma\right)$$ conditional independence

$$\phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right) = \int_{\mathbb{R}^{n_{j_1}}} \left[\int_{\mathbb{R}} p_{\mathbf{Y^{j_1}}|X_{j_1}}(\mathbf{y};x) p_{X_{j_1}|X_j}(x;X_j)\,\mathrm{d}\,x\right] e^{i\mathbf{y}^T\mathbf{u}^{j_1}}\,\mathrm{d}\,\mathbf{y}$$

Characteristic Function : Pre-order Tree Traversal

Characteristic Function:

$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \exp\left(iX_j \sum_{k \in \text{desTips}(j)} u_k - \sigma \sum_{k \in \text{des}\, j} t_k \left| \sum_{l \in \text{desTips}(k)} u_l \right| \right)$$



Propagation:

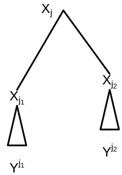
$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right) \times \phi_{\mathbf{Y^{j_2}}|X_j}\left(\mathbf{u}^{j_2};\sigma\right)$$ conditional independence

$$\begin{aligned}
\phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right) &= \int_{\mathbb{R}^{n_{j_1}}} \left[\int_{\mathbb{R}} p_{\mathbf{Y^{j_1}}|X_{j_1}}(\mathbf{y};x) p_{X_{j_1}|X_j}(x;X_j)\, \mathrm{d}\,x \right] e^{i\mathbf{y}^T\mathbf{u}^{j_1}}\, \mathrm{d}\,\mathbf{y} \\
&= \int_{\mathbb{R}} \left[\int_{\mathbb{R}^{n_{j_1}}} p_{\mathbf{Y^{j_1}}|X_{j_1}}(\mathbf{y};x) e^{i\mathbf{y}^T\mathbf{u}^{j_1}}\, \mathrm{d}\,\mathbf{y} \right] p_{X_{j_1}|X_j}(x;X_j)\, \mathrm{d}\,x
\end{aligned}$$

Characteristic Function : Pre-order Tree Traversal

Characteristic Function:

$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \exp\left(iX_j \sum_{k\in \mathrm{desTips}(j)} u_k - \sigma \sum_{k\in \mathrm{des}\, j} t_k \left| \sum_{l\in \mathrm{desTips}(k)} u_l \right| \right)$$



Propagation:

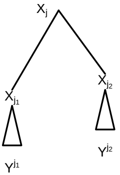
$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right)\times\phi_{\mathbf{Y^{j_2}}|X_j}\left(\mathbf{u}^{j_2};\sigma\right) \quad \text{conditional independence}$$

$$\begin{aligned}
\phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right) &= \int_{\mathbb{R}^{n_{j_1}}} \left[\int_{\mathbb{R}} p_{\mathbf{Y^{j_1}}|X_{j_1}}(\mathbf{y};x) p_{X_{j_1}|X_j}(x;X_j)\,\mathrm{d}\,x\right] e^{i\mathbf{y}^T\mathbf{u}^{j_1}}\,\mathrm{d}\,\mathbf{y} \\
&= \int_{\mathbb{R}} \left[\int_{\mathbb{R}^{n_{j_1}}} p_{\mathbf{Y^{j_1}}|X_{j_1}}(\mathbf{y};x) e^{i\mathbf{y}^T\mathbf{u}^{j_1}}\,\mathrm{d}\,\mathbf{y}\right] p_{X_{j_1}|X_j}(x;X_j)\,\mathrm{d}\,x \\
&= \int_{\mathbb{R}} \left[\phi_{\mathbf{Y^{j_1}}|X_{j_1}=x}\left(\mathbf{u}^{j_1};\sigma\right)\right] p_{X_{j_1}|X_j}(x;X_j)\,\mathrm{d}\,x
\end{aligned}$$

Characteristic Function : Pre-order Tree Traversal

Characteristic Function:

$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \exp\left(iX_j \sum_{k\in \text{desTips}(j)} u_k - \sigma \sum_{k\in \text{des}\, j} t_k \left|\sum_{l\in \text{desTips}(k)} u_l\right|\right)$$



Propagation:

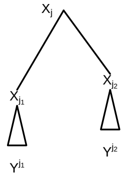
$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right)\times\phi_{\mathbf{Y^{j_2}}|X_j}\left(\mathbf{u}^{j_2};\sigma\right)$$ conditional independence

$$\phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right) = \exp\left(-\sigma \sum_{k\in \text{des}\, j_1} t_k \left|\sum_{l\in \text{desTips}(k)} u_l\right|\right)\times \mathbb{E}\left[\exp\left(iZ \sum_{k\in \text{desTips}(j_1)} u_k\right)\right] \text{ with } Z \sim \mathcal{C}(X_j, \sigma t_{j_1})$$

Characteristic Function : Pre-order Tree Traversal

Characteristic Function:

$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \exp\left(iX_j \sum_{k\in \mathrm{desTips}(j)} u_k - \sigma \sum_{k\in \mathrm{des}\, j} t_k \left|\sum_{l\in \mathrm{desTips}(k)} u_l\right|\right)$$



Propagation:

$$\phi_{\mathbf{Y^j}|X_j}\left(\mathbf{u}^j;\sigma\right) = \phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right)\times\phi_{\mathbf{Y^{j_2}}|X_j}\left(\mathbf{u}^{j_2};\sigma\right)$$ conditional independence

$$\phi_{\mathbf{Y^{j_1}}|X_j}\left(\mathbf{u}^{j_1};\sigma\right) = \exp\left(-\sigma \sum_{k\in \mathrm{des}\, j_1} t_k \left|\sum_{l\in \mathrm{desTips}(k)} u_l\right|\right)\times$$

$$\exp\left(iX_j \sum_{k\in \mathrm{desTips}(j_1)} u_k - \sigma t_{j_1}\left|\sum_{k\in \mathrm{desTips}(j_1)} u_k\right|\right)$$

back

Likelihood

$$p(\mathbf{Y} \mid \mu, \sigma) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\mathbf{u}^T\mathbf{Y}} \phi_{\mathbf{Y}|X_{\text{root}}}(\mathbf{u}; \sigma) \,\mathrm{d}\,\mathbf{u}$$

$$p(\mathbf{Y} \mid \mu, \sigma) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{-i\mathbf{u}^T\mathbf{Y}} \exp\left(i\mu \sum_{k=1}^{n} u_k - \sigma \sum_{j \neq \text{root}} t_j \left| \sum_{k \in \text{desTips}(j)} u_k \right| \right) \mathrm{d}\,\mathbf{u}$$

Idea:

- take a node j with two descending tips k and l
- change of variable : $v_k = u_k + u_l$ and $v_l = u_l$
- integrate out v_l.

+

back

Likelihood Computation

Partial Likelihood:

$$p\left(\mathbf{y}^r \mid z_r, \sigma, \mathcal{T}\right) = \frac{1}{(2\pi)^{|\mathrm{L}_\mathcal{T}|}} \int_{-\infty}^{+\infty} \sum_{b \in \mathrm{L}_\mathcal{T}} C_{r,b}^{\mathrm{sgn}(u)} \exp\left(-\sigma t_{r:b}|u| - iu(y_b - z_r)\right) du.$$

Likelihood Computation

Partial Likelihood:

$$p\left(\mathbf{y}^r \mid z_r, \sigma, \mathcal{T}\right) = \frac{1}{(2\pi)^{|\mathrm{L}_{\mathcal{T}}|}} \int_{-\infty}^{+\infty} \sum_{b \in \mathrm{L}_{\mathcal{T}}} C_{r,b}^{\mathrm{sgn}(u)} \exp\left(-\sigma t_{r:b}|u| - iu(y_b - z_r)\right) du.$$

Recursion formula:

$$C_{j,b} = C_{m,b} \sum_{c \in \mathrm{desTips}(k)} \left(\frac{C_{k,c}}{\sigma(t_{j:c} - t_{j:b}) + i(y_c - y_b)} + \frac{\overline{C_{k,c}}}{\sigma(t_{j:b} + t_{j:c}) - i(y_c - y_b)} \right)$$

Likelihood Computation

Partial Likelihood:

$$p\left(\mathbf{y}^r \mid z_r, \sigma, \mathcal{T}\right) = \frac{1}{(2\pi)^{|\mathrm{L}_\mathcal{T}|}} \int_{-\infty}^{+\infty} \sum_{b \in \mathrm{L}_\mathcal{T}} C_{r,b}^{\mathrm{sgn}(u)} \exp\left(-\sigma t_{r:b}|u| - iu(y_b - z_r)\right) du.$$

Recursion formula:

$$C_{j,b} = C_{m,b} \sum_{c \in \mathrm{desTips}(k)} \left(\frac{C_{k,c}}{\sigma(t_{j:c} - t_{j:b}) + i(y_c - y_b)} + \frac{\overline{C_{k,c}}}{\sigma(t_{j:b} + t_{j:c}) - i(y_c - y_b)} \right)$$

Initialization for any tip i: $C_{i,i}^{+} = C_{i,i}^{-} = 1$.

Likelihood Computation

Conditional Independence:

$$p\left(\mathbf{y}^k \;\middle|\; z_k, \sigma, \mathcal{T}_k^-\right) = p\left(\mathbf{y}^i \;\middle|\; z_k, \sigma, \mathcal{T}_i\right) p\left(\mathbf{y}^j \;\middle|\; z_k, \sigma, \mathcal{T}_j\right)$$

Likelihood Computation

Conditional Independence:

$$p\left(\mathbf{y}^k \,\middle|\, z_k, \sigma, \mathcal{T}_k^-\right) = p\left(\mathbf{y}^i \,\middle|\, z_k, \sigma, \mathcal{T}_i\right) p\left(\mathbf{y}^j \,\middle|\, z_k, \sigma, \mathcal{T}_j\right)$$

Integration on parent branch:

$$p\left(\mathbf{y}^k \,\middle|\, z_{\mathrm{pa}(k)}, \sigma, \mathcal{T}_k^-\right) = \int_{-\infty}^{+\infty} p\left(\mathbf{y}^k \,\middle|\, z_k, \sigma, \mathcal{T}_k^-\right) p\left(z_k \,\middle|\, z_{\mathrm{pa}(k)}, \sigma\right) dz_k$$

Likelihood Computation

Conditional Independence:

$$p\left(\mathbf{y}^k \;\middle|\; z_k, \sigma, \mathcal{T}_k^-\right) = p\left(\mathbf{y}^i \;\middle|\; z_k, \sigma, \mathcal{T}_i\right) p\left(\mathbf{y}^j \;\middle|\; z_k, \sigma, \mathcal{T}_j\right)$$

Integration on parent branch:

$$p\left(\mathbf{y}^k \;\middle|\; z_{\mathrm{pa}(k)}, \sigma, \mathcal{T}_k^-\right) = \int_{-\infty}^{+\infty} p\left(\mathbf{y}^k \;\middle|\; z_k, \sigma, \mathcal{T}_k^-\right) p\left(z_k \;\middle|\; z_{\mathrm{pa}(k)}, \sigma\right) dz_k$$

Recursion:

$$\begin{aligned} p\left(\mathbf{y}^k \;\middle|\; z_{\mathrm{pa}(k)}, \sigma, \mathcal{T}_k^-\right) = \frac{1}{(2\pi)^{|\mathrm{L}_k|}} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} \exp\left(i(u+v)z_{\mathrm{pa}(k)} - \sigma t_k|u+v|\right) \\ \times \left[\sum_{b\in\mathrm{L}_i} C_{k,b}^{\mathrm{sgn}(u)} \exp\left(-\sigma t_{k:b}|u| - iuy_b\right)\right]\left[\sum_{c\in\mathrm{L}_j} C_{k,c}^{\mathrm{sgn}(v)} \exp\left(-\sigma t_{k:c}|v| - iuy_c\right)\right] dudv \end{aligned}$$

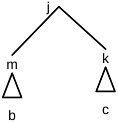
Likelihood Computation

Root integration:

$$p(\mathbf{Y} \mid \mu = 0, \sigma) = \frac{2}{(2\pi)^n} \sum_{b \in \mathrm{desTips}(r)} \frac{R_{r,b}\sigma t_{r:b} + I_{r,b} y_b}{(\sigma t_{r:b})^2 + y_b^2}$$

back

Likelihood Computation Algorithm



Recursion Formula:

$$C_{j,b} = C_{m,b} \sum_{c \in \mathrm{desTips}(k)} \left(\frac{C_{k,c}}{\sigma(t_{j:c} - t_{j:b}) + i(y_c - y_b)} + \frac{\overline{C_{k,c}}}{\sigma(t_{j:b} + t_{j:c}) - i(y_c - y_b)} \right)$$

Exact:

Can compute the integral explicitly, with one traversal of the tree.

Complexity:

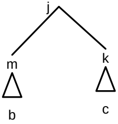
Quadratic in the number of tips.

Stability:

Sums of large positive and negative numbers: numerical issues.

Theoretical problem:

Likelihood Computation Algorithm



Recursion Formula:

$$C_{j,b} = C_{m,b} \sum_{c \in \mathrm{desTips}(k)} \left(\frac{C_{k,c}}{\sigma(t_{j:c} - t_{j:b}) + i(y_c - y_b)} + \frac{\overline{C_{k,c}}}{\sigma(t_{j:b} + t_{j:c}) - i(y_c - y_b)} \right)$$

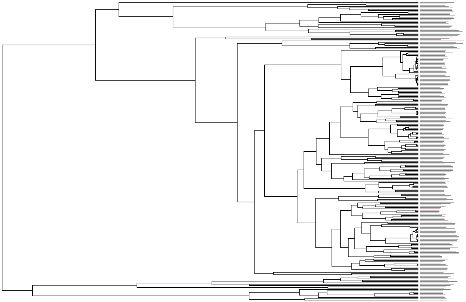
Exact:
Can compute the integral explicitly, with one traversal of the tree.

Complexity:
Quadratic in the number of tips.

Stability:
Sums of large positive and negative numbers: numerical issues.

Theoretical problem: Division by zero !

Chelonia Dataset



Jaffe et al. (2011)

Dermochelys Coriacea

Homopus Aerolatus

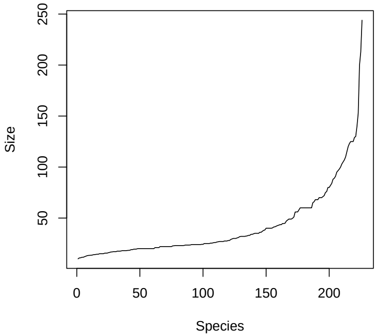
```
summary(data)

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   2.303   2.996   3.296   3.482   3.892   5.497
```

Chelonia Dataset

```
summary(exp(data))

##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   10.00   20.00   27.00   41.67   49.00  244.00
```

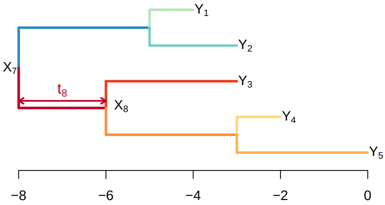
Chelonia Dataset

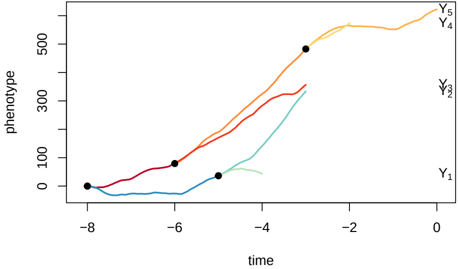
```
summary(exp(data))

##     Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    10.00   20.00   27.00   41.67   49.00  244.00
```



back

Integrated OU





Integrated Ornstein-Uhlenbeck

$$\begin{pmatrix} V_i \\ X_i \end{pmatrix} \Big| \begin{pmatrix} V_{\mathrm{pa}(i)} \\ X_{\mathrm{pa}(i)} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} e^{-\alpha t_i} & 0 \\ (1-e^{-\alpha t_i})/\alpha & 1 \end{pmatrix} \begin{pmatrix} V_{\mathrm{pa}(i)} \\ X_{\mathrm{pa}(i)} \end{pmatrix} + \begin{pmatrix} (1-e^{-\alpha t_i})\beta \\ (t_i - (1-e^{-\alpha t_i})/\alpha)\beta \end{pmatrix}; \frac{\sigma^2}{2\alpha} \begin{pmatrix} (1-e^{-2\alpha t_i}) & \frac{(1-e^{-\alpha t_i})^2}{\alpha} \\ \frac{(1-e^{-\alpha t_i})^2}{\alpha} & \frac{2t_i}{\alpha} - 4\frac{(1-e^{-\alpha t_i})}{\alpha^2} + \frac{(1-e^{-2\alpha t_i})}{\alpha^2} \end{pmatrix} \right)$$

Existing Approaches

RRW: Lemey et al. (2010); Fisher et al. (2021)

$$p(\phi, \sigma \mid \mathbf{Y}) \propto p(\mathbf{Y} \mid \phi, \sigma)\, p(\phi)\, p(\sigma) \quad \text{with } \phi_i \sim \text{Inv-Gam}(1/2, 1/2)$$

General Stable Process: Elliot and Mooers (2014)

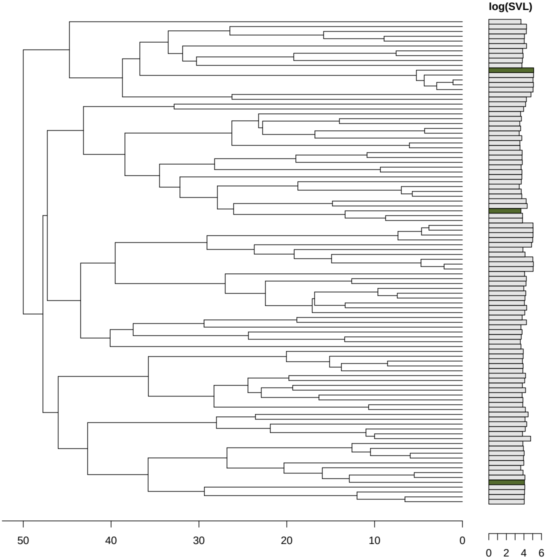
General Lévy Process: Landis et al. (2013); Duchen et al. (2017); Landis and Schraiber (2017)

$\rightarrow$ large latent space, numerical integration

New approach: Direct numerical likelihood maximization.

Greater Antilles Anolis Lizards

(Mahler et al., 2013)



Anolis equestris

Anolis porcatus

Anolis sagrei

Lizard Dataset

```
library(cauphy)
```

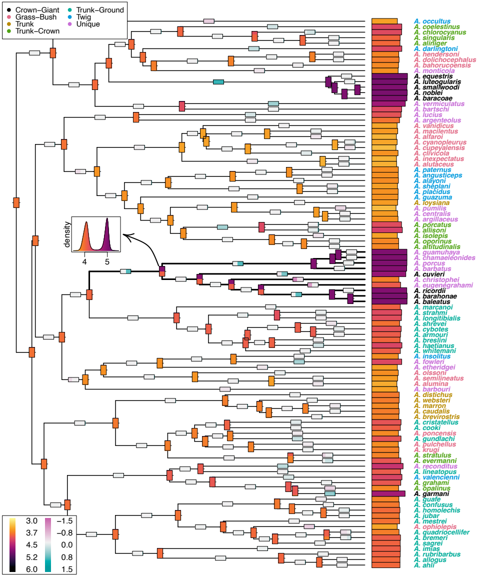
```
fitContinuous(phy, svl, model = "BM")$opt$lnL

## [1] -4.700404
```

```
fitContinuous(phy, svl, model = "lambda")$opt$lnL

## [1] -4.700404
```

```
fitCauchy(phy, svl, model = "cauchy", method = "fixed.root")$logLik

## [1] 4.441921
```

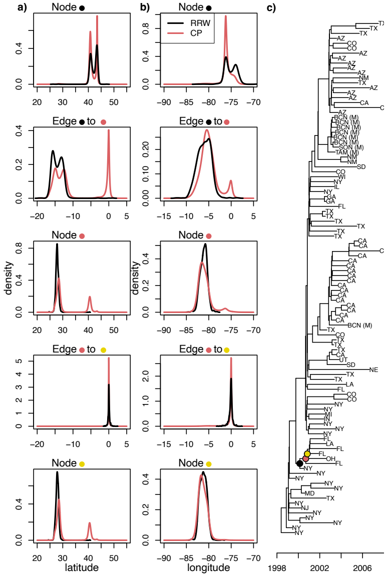
```
fitCauchy(phy, svl, model = "lambda", method = "fixed.root")$logLik

## [1] 4.926054
```

Lizard Dataset



West Nile Virus

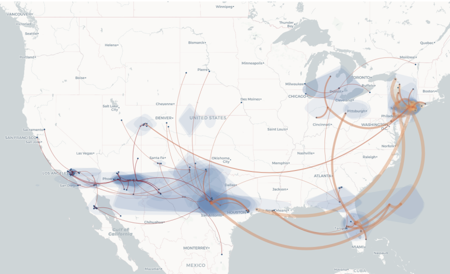
(Pybus et al., 2012)



WNV using Evolaps

(Pybus et al., 2012; Chevenet et al., 2024)

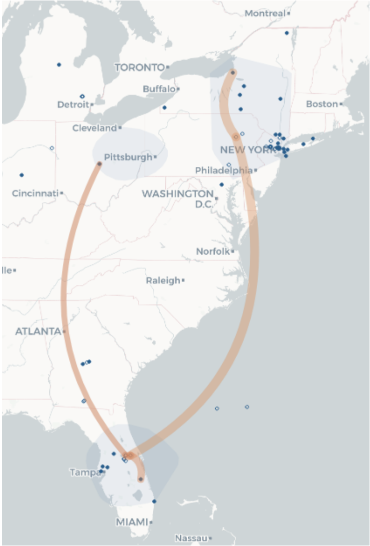
CP:



Cauchy RRW:

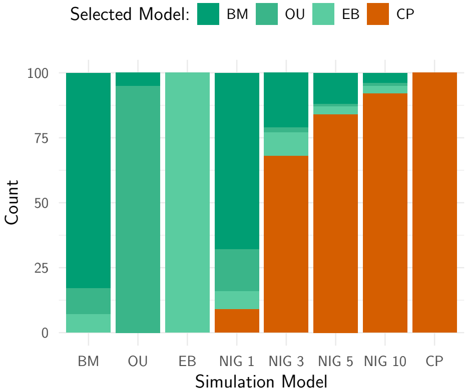
WNV using Evolaps (Pybus et al., 2012; Chevenet et al., 2024)

CP:



Cauchy RRW:

Simulation Study - Model Selection

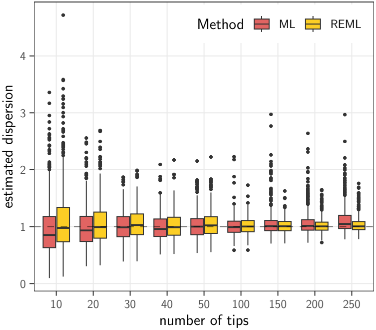


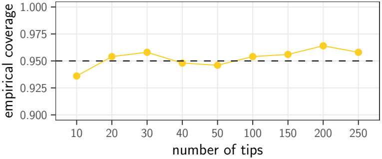
Simulations:

Design similar to Landis and Schraiber (2017).

Model selection with AIC criteria.

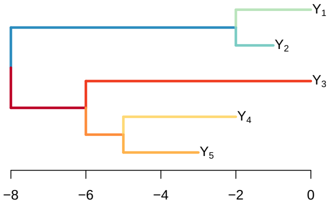
Simulation Study - Dispersion Estimation

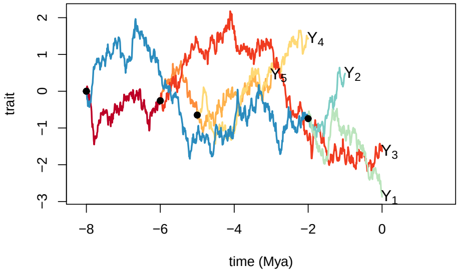




REML: Re-root at a tip.

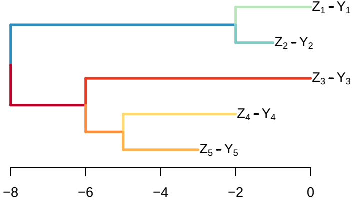
Measurement Error

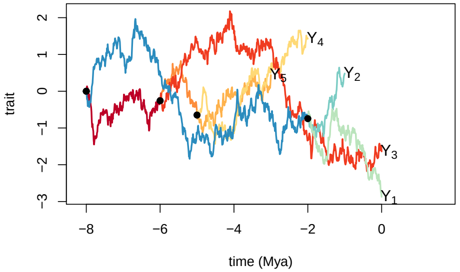




- Heredity: $X_8 | X_7 \sim \mathcal{N}(X_7, \sigma^2 t_8)$
- Covariances: $\mathbb{C}\text{ov}(Y_i, Y_j) = \sigma^2 V_{ij}$
- Distribution: $\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma^2 \mathbf{V})$

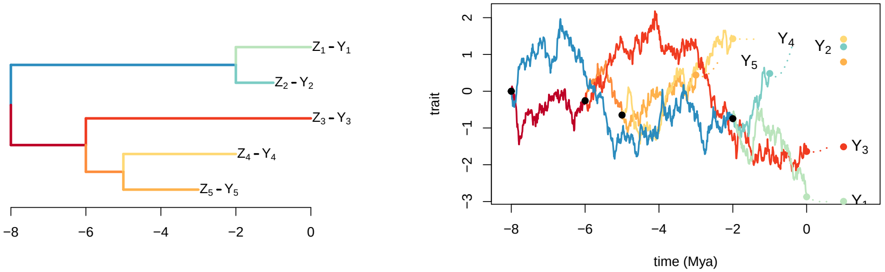
Measurement Error





- Heredity: $X_8 | X_7 \sim \mathcal{N}(X_7, \sigma^2 t_8)$
- Covariances: $\mathbb{C}\text{ov}(Y_i, Y_j) = \sigma^2 V_{ij}$
- Distribution: $\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma^2 \mathbf{V})$

Measurement Error



- Observation: $Y_1 | Z_1 \sim \mathcal{N}(Z_1, s^2)$
- Covariances: $\mathbb{C}\text{ov}(Y_i, Y_j) = \sigma^2 V_{ij} + \sigma_e^2 \delta_{ij}$
- Distribution: $\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma^2 \mathbf{V} + s^2 \mathbf{I}_n)$

Pagel's Lambda (Pagel, 1999)

Relax the BM variance structure:

$$\mathbf{V}(\lambda)_{ii} = \mathbf{V}_{ii}$$
$$\mathbf{V}(\lambda)_{ij} = \lambda\mathbf{V}_{ij}$$

Pagel's Lambda

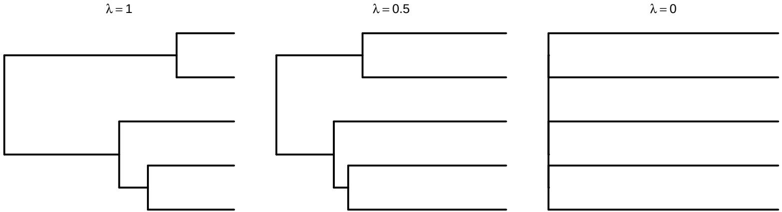
(Pagel, 1999)

Relax the BM variance structure:

$$\mathbf{V}(\lambda)_{ii} = \mathbf{V}_{ii}$$
$$\mathbf{V}(\lambda)_{ij} = \lambda \mathbf{V}_{ij}$$

Equivalent to running a BM on a modified tree with:

$$t(\lambda)_i = \begin{cases} \lambda t_i & \text{if } i \text{ internal node} \\ \lambda t_i + (1-\lambda) T_i & \text{if } i \text{ leaf} \end{cases}$$



Measurement Error and Pagel's λ (Leventhal and Bonhoeffer, 2016)

Pagel's Lambda: (ultrametric tree with height T)

$$\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma_\lambda^2 \mathbf{V}(\lambda)) \qquad \mathbf{V}(\lambda) = \lambda \mathbf{V} + (1 - \lambda) T \mathbf{I}$$

Measurement Error and Pagel's λ (Leventhal and Bonhoeffer, 2016)

Pagel's Lambda: (ultrametric tree with height T)

$$\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma_\lambda^2 \mathbf{V}(\lambda)) \qquad \mathbf{V}(\lambda) = \lambda \mathbf{V} + (1 - \lambda) T \mathbf{I}$$

BM with measurement error:

$$\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma^2 \mathbf{V} + s^2 \mathbf{I})$$

Measurement Error and Pagel's λ (Leventhal and Bonhoeffer, 2016)

Pagel's Lambda: (ultrametric tree with height T)

$$\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma_\lambda^2 \mathbf{V}(\lambda)) \qquad \mathbf{V}(\lambda) = \lambda \mathbf{V} + (1 - \lambda) T \mathbf{I}$$

BM with measurement error:

$$\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma^2 \mathbf{V} + s^2 \mathbf{I})$$

Equivalent if:

$$\lambda = \frac{\sigma^2 T}{\sigma^2 T + s^2} \qquad \sigma_\lambda^2 = \sigma^2 + s^2 / T$$

Measurement Error and Pagel's λ (Leventhal and Bonhoeffer, 2016)

Pagel's Lambda: (ultrametric tree with height T)

$$\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma_\lambda^2 \mathbf{V}(\lambda)) \qquad \mathbf{V}(\lambda) = \lambda \mathbf{V} + (1-\lambda) T \mathbf{I}$$

BM with measurement error:

$$\mathbf{Y} \sim \mathcal{N}(\mu \mathbf{1}_n, \sigma^2 \mathbf{V} + s^2 \mathbf{I})$$

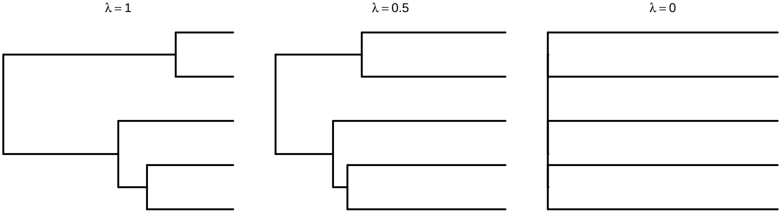
Equivalent if:

$$\lambda = \frac{\sigma^2 T}{\sigma^2 T + s^2} \qquad \sigma_\lambda^2 = \sigma^2 + s^2 / T$$

λ is the phylogenetic heritability

Cauchy Pagel's λ

$$t(\lambda)_i = \begin{cases} \lambda t_i & \text{if } i \text{ internal node} \\ \lambda t_i + (1-\lambda)T & \text{if } i \text{ leaf} \end{cases}$$



Cauchy Pagel's λ

$$t(\lambda)_i = \begin{cases} \lambda t_i & \text{if } i \text{ internal node} \\ \lambda t_i + (1-\lambda) T & \text{if } i \text{ leaf} \end{cases}$$

Cauchy on the transformed tree:

$$\log \phi_{\mathbf{Y}|X_{\text{root}}}(\mathbf{u}; \sigma, \lambda) = i\mu \sum_{k=1}^{n} u_k - \sigma \sum_{j \neq \text{root}} t(\lambda)_j \left| \sum_{k \in \text{desTips}(j)} u_k \right|$$

Cauchy Pagel's λ

$$t(\lambda)_i = \begin{cases} \lambda t_i & \text{if } i \text{ internal node} \\ \lambda t_i + (1-\lambda)T & \text{if } i \text{ leaf} \end{cases}$$

Cauchy on the transformed tree:

$$\log \phi_{\mathbf{Y}|X_{\text{root}}}(\mathbf{u}; \sigma, \lambda) = i\mu \sum_{k=1}^{n} u_k - \sigma_\lambda \lambda \sum_{j \neq \text{root}} t_j \left| \sum_{k \in \text{desTips}(j)} u_k \right| - \sigma_\lambda (1-\lambda) T \sum_{i \ tip} |u_i|$$

Cauchy Pagel's λ

$$t(\lambda)_i = \begin{cases} \lambda t_i & \text{if } i \text{ internal node} \\ \lambda t_i + (1-\lambda)T & \text{if } i \text{ leaf} \end{cases}$$

Cauchy on the transformed tree:

$$\log \phi_{\mathbf{Y}|X_{\text{root}}}(\mathbf{u}; \sigma, \lambda) = i\mu \sum_{k=1}^{n} u_k - \sigma_\lambda \lambda \sum_{j \neq \text{root}} t_j \left| \sum_{k \in \text{desTips}(j)} u_k \right| - \sigma_\lambda (1-\lambda) T \sum_{i \ tip} |u_i|$$

Cauchy Errors:

$$\sigma_\lambda \lambda = \sigma \qquad \sigma_\lambda (1-\lambda) T = s$$

Cauchy Pagel's λ

$$t(\lambda)_i = \begin{cases} \lambda t_i & \text{if } i \text{ internal node} \\ \lambda t_i + (1-\lambda)T & \text{if } i \text{ leaf} \end{cases}$$

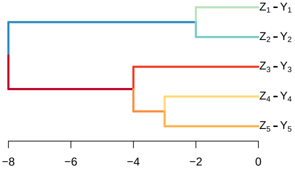
Cauchy on the transformed tree:

$$\log \phi_{\mathbf{Y}|X_{\text{root}}}(\mathbf{u}; \sigma, \lambda) = i\mu \sum_{k=1}^{n} u_k - \sigma_\lambda \lambda \sum_{j \neq \text{root}} t_j \left| \sum_{k \in \text{desTips}(j)} u_k \right| - \sigma_\lambda (1-\lambda) T \sum_{i \ tip} |u_i|$$

Cauchy Errors:

$$\lambda = \frac{\sigma T}{\sigma T + s} \qquad \sigma_\lambda = \sigma + s/T$$

Cauchy Errors



- Heredity: $X_8 | X_7 \sim \mathcal{C}(X_7, \sigma t_8)$
- Observation: $Y_1 | Z_1 \sim \mathcal{C}(Z_1, s)$
- Marginal: $Y_i \sim \mathcal{C}(\mu, \sigma T + s)$

Equivalent to Pagel's λ.

Pagel's λ tree transform

Easy to implement:
Transform tree for each λ.

Phylogenetic Heritability:

- $\lambda = 0$: no heritability.
- $\lambda = 1$: no individual variation.

Individual variation:
Measurement error, intra-specific variation, multiple measurements, ...

Valid for any α-stable process.

Multivariate α Stable Distribution (Nolan, 2005)

Definition:

$$\mathbf{X} \sim \mathcal{MC}_p \iff \mathbf{a}^T \mathbf{X} \sim \mathcal{C} \quad , \ \forall \mathbf{a} \in \mathbb{R}^p.$$

Characterization:

$$\mathbf{X} \sim \mathcal{MC}_p \iff \phi_{\mathbf{X}}(\mathbf{u}) = \mathbb{E}\left[e^{\mathbf{u}^T \mathbf{X}}\right] = e^{i\mathbf{u}^T \boldsymbol{\mu} - \gamma(\mathbf{u})}$$

with:

$$\gamma(\mathbf{u}) = \int_{\mathcal{S}_p} \left|\mathbf{u}^T \mathbf{s}\right| \mathrm{d}\,\sigma(\mathbf{s})$$

and $\sigma(\mathbf{s})$ is a *spectral measure* on the sphere $\mathcal{S}_p$.

Special Case: Linear Combinations (Kidmose, 2001)

Assumption:

$$\mathbf{X} = \mathbf{A}\mathbf{V} \quad \text{with} \quad \mathbf{A} : \ p \times q \quad \text{and} \quad \mathbf{V} \text{ vector of } q \text{ iid Cauchy}$$

Special Case: Linear Combinations (Kidmose, 2001)

Assumption:

$$\mathbf{X} = \mathbf{AV} \quad \text{with} \quad \mathbf{A}: \ p \times q \quad \text{and} \quad \mathbf{V} \text{ vector of } q \text{ iid Cauchy}$$

Notes:

- If $p = q$, then getting the density is easy.
- $q > p$ is allowed.
- Cauchy on a tree : special case with $p = n$ and $q = 2n - 2$.

Special Case: Linear Combinations (Kidmose, 2001)

Assumption:

$$\mathbf{X} = \mathbf{A}\mathbf{V} \quad \text{with} \quad \mathbf{A} : \; p \times q \quad \text{and} \quad \mathbf{V} \text{ vector of } q \text{ iid Cauchy}$$

Notes:

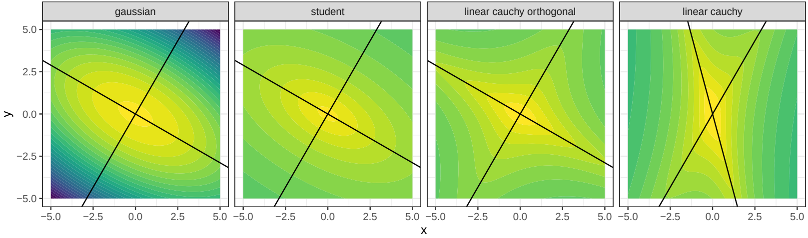
- If $p = q$, then getting the density is easy.
- $q > p$ is allowed.
- Cauchy on a tree : special case with $p = n$ and $q = 2n - 2$.

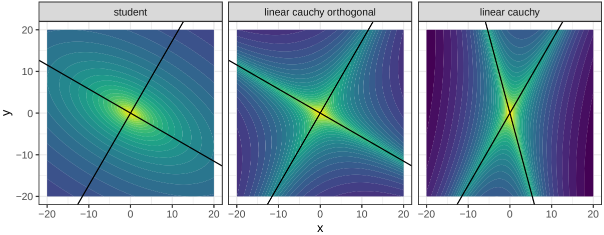
Characteristic Function: Assuming $V_i \sim \mathcal{C}(\mu_i, \sigma_i)$

$$\phi_{\mathbf{X}}(\mathbf{u}) = e^{i\mathbf{u}^T\boldsymbol{\mu} - \gamma(\mathbf{u})} \quad \text{with} \quad \gamma(\mathbf{u}) = \int_{\mathcal{S}_p} \left|\mathbf{u}^T\mathbf{s}\right| \mathrm{d}\,\sigma(\mathbf{s}) \quad \text{and}$$

$$\sigma(\mathbf{s}) = \sum_{i=1}^{q} \frac{1}{2}\sigma_i \sqrt{\mathbf{A}_i^T\mathbf{A}_i}(\delta(\mathbf{s} - \mathbf{s}_i) + \delta(\mathbf{s} + \mathbf{s}_i)).$$

Case $p = q = 2$


gaussian
student
linear cauchy orthogonal
linear cauchy
y
x
5.0
2.5
0.0
−2.5
−5.0


student
linear cauchy orthogonal
linear cauchy
y
x
20
10
0
−10
−20

Special Case: Linear Combinations (Kidmose, 2001)

Assumption:

$$\mathbf{X} = \mathbf{AV} \quad \text{with} \quad \mathbf{A} : \ p \times q \quad \text{and} \quad \mathbf{V} \text{ vector of } q \text{ iid Cauchy}$$

Notes:

- If $p = q$, then getting the density is easy.
- $q > p$ is allowed.
- Cauchy on a tree : special case with $p = n$ and $q = 2n - 2$.

Perspectives:

- For $p = 2$, and any q: we can get the density.
- It looks like a mixture of Cauchy-Like distributions.
- Can we do any p ?