Cosmological time delay estimation with Continuous Auto-Regressive Moving Average processes

Antoine D. Meyer (antoine.meyer@cfa.harvard.edu - adm18@ic.ac.uk)
Statistics Section, Department of Mathematics, Imperial College London
Predoc at CfA
In collaboration with David van Dyk, Hyungsuk Tak, Aneta Siemiginowska

CHASC
19 April 2022

The gravitational lensing phenomenon

- **Light rays are bent** by the strong gravitational field of the intervening object (i.e. the lens) on their way to Earth.
- Different rays travel along different paths of different distances, and therefore **arrive to the observer at different times** (separated by a delay Δ).

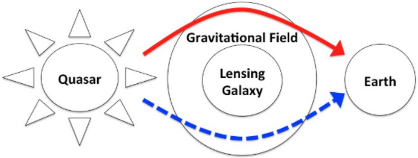


Figure: Source: Tak et al. [2017]

- Multiple copies of the original light curve, **brightness fluctuations are observed at different times** in the different copies.
- Refsdal [1964]: Estimates of time delays can be used to **constrain important cosmological parameters** such as H_0.

A challenging statistical problem

Goal: estimate the time shift (x-axis) between multiple lensed light curves

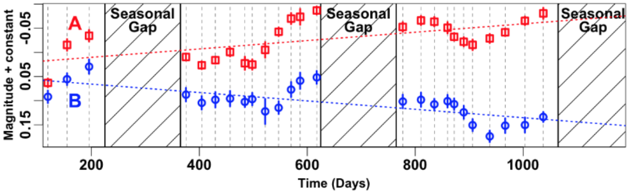


Figure: Source: Tak et al. [2017]

Data is subject to:

- Irregular sampling (observational patterns)
- Seasonal gaps (celestial cycles)
- Brightness magnifications due to multiple effects (strong lensing, micro-lensing)
- Measurement errors (heteroskedastic)

Current methods:

Two families of methods for time delay estimation:

- Grid-based optimization methods (≈ non-parametric)
  - Minimize measure of distance between light curves, on a grid of Δ values
  - Produce uncertainty estimates using Monte Carlo simulation
  - Computationally expensive!

- Statistics modelling the stochastic variability og AGN light curves
  - More principled
  - "Direct" quantification of uncertainty in time delay parameter

Time Delay Challenge (TDC, Liao et al. [2015]) surveyed and evaluated a variety of time delay estimation techniques

- Winners: COSMOGRAIL collaboration (combine estimates from 4 different non-parametric techniques)
- Tak et al. [2017] did very well too! Bayesian method + DRW process to model AGN variability)

Goals of this project

Goals of this project :

- Improve Bayesian time delay estimation method, based on Tak et al. [2017]
- Improve on the accuracy and applicability of time delay estimation method
- Address some of the computational limitations of current estimation strategies.

Current limitations:

- (Tak et al. [2017]) limited applicability of the DRW process to model AGN light curves.
- Difficult computation, in part due to multi-modality issues
- High sensitivity to initial guess for Δ, most methods require method to compute plausible/starting value of Δ.

Our contributions

Limitation: use of the DRW to model AGN light curves restricts the range of observations on which time delay method can be applied

- Recent success in modeling AGN light curves with flexible CARMA(p, q) in astrophysics literature [Kelly et al., 2014, Moreno et al., 2019].
- AGNs and particularly quasars are sources for which strong gravitational lensing is more likely to happen
- Generalization of DRW, which is a CARMA(1,0).

Main development: update the intrinsic light curve model from DRW process to flexible **Continuous Auto-Regressive Moving Average (CARMA) process**

- Finer modeling tool can better fit a wider range of observations.
- Better fit to the lightcurve data $\rightarrow$ better accuracy

Our method: TD-CARMA
Tak et al. [2017]: TD-DRW

Contributions

Additional development: update the parameter inference algorithm from MCMC [Tak et al., 2017] to nested sampling (MultiNest)

- Deal with **multi-modality** of CARMA(p, q) and time delay parameters
  - MultiNest identifies multiple modes in posterior distribution
  - MultiNest output can be used to quantify relative probability of the modes
- **Blind Search:** no initial value of Δ required
  - Most existing time delay methods are highly sensitive to initial values
  - Require an extra method to compute a plausible initial value
  - Tak et al. [2017] compute the expensive profile likelihood
- **Model Selection:** MultiNest estimates the Bayesian evidence (marginal distribution of the data) at no extra computational cost

Data

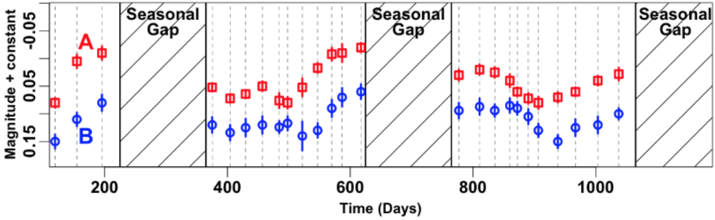


Figure: Source: Tak et al. [2017]

$$\boxed{\mathbf{D} = \{t_i, x_i, \delta_i^x, y_i, \delta_i^y\}_{i=1}^n}$$

- Observation times $\mathbf{t} = \{t_1, \ldots, t_n\}$
- Observed magnitudes $\mathbf{x} = \{x_1, \ldots, x_n\}$, and $\mathbf{y}$
- Measurement errors $\delta^{\mathbf{x}} = \{\delta_1^x, \ldots, \delta_n^x\}$ and $\delta^{\mathbf{y}}$ (standard deviation).

Time Delay Estimation Framework

Assumptions of the time delay model:

- Assumption I: $\boldsymbol{x}$ and $\boldsymbol{y}$ = discrete realizations of unobserved continuous lightcurves $x(t)$ and $y(t)$ (true source magnitudes), $t \in \mathbb{R}$.
- Assumption II: $y(t)$ is time and magnitude shifted version of $x(t)$.

Let's translate them into our model:

1. Strong lensing effect 1: Time shift

$$y(t) = x(t - \Delta) \tag{1}$$

2. Strong lensing effect 2: different average magnitudes

$$y(t) = x(t - \Delta) + \theta_0 \tag{2}$$

3. Micro lensing effect: extrinsic long-term variability

$$y(t) = x(t - \Delta) + \boldsymbol{w}_m(t - \Delta)\boldsymbol{\theta} \tag{3}$$

- $\boldsymbol{w}_m(t - \Delta) := \{1, t - \Delta, \ldots, (t - \Delta)^m\}$ vector of polynomial time variables
- $\boldsymbol{\theta} = \{\theta_0, \ldots, \theta_m\}$ micro-lensing coefficients

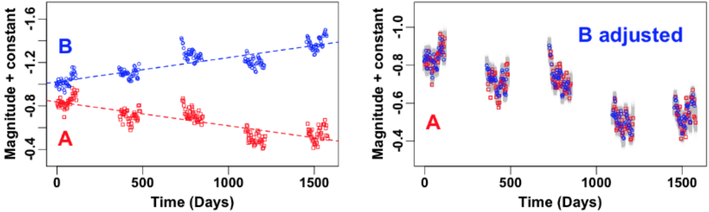
Time Delay Estimation Framework

Plan:

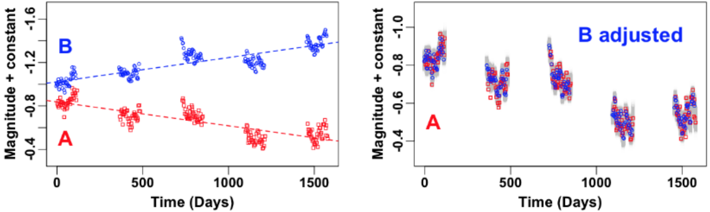
1. **Reconstruct intrinsic light curve** from its lensed counterparts
2. Model stochastic variability in intrinsic lightcurve with CARMA process

Reconstructing the intrinsic light curve:

- Given Δ and θ, construct measurements for the intrinsic light curve, denoted $\boldsymbol{z} = \{z_j\}_{j=1}^{2n}$.
- Measured at times $\boldsymbol{t}^{\boldsymbol{\Delta}} = \{t_i\}_{i=1}^{n} \cup \{t_i - \Delta\}_{i=1}^{n}$.



Time Delay Estimation Framework



$$z_j = \begin{cases} x_i & \text{for some } i \text{ if } t_j^{\Delta} \text{ is in } \boldsymbol{t}, \\ y_i - \boldsymbol{w}_m(t_j - \Delta)\boldsymbol{\theta} & \text{for some } i \text{ if } t_j^{\Delta} \text{ is in } \boldsymbol{t} - \Delta, \end{cases} \quad (4)$$

- Similarly, for the vector measurement error standard deviations $\{\delta_i^z\}_{i=1}^{2n}$:

$$\delta_j^z = \begin{cases} \delta_i^x & \text{for some } i \text{ if } t_j^{\Delta} \text{ is in } \boldsymbol{t}, \\ \delta_i^y & \text{for some } i \text{ if } t_j^{\Delta} \text{ is in } \boldsymbol{t} - \Delta. \end{cases} \quad (5)$$

Time Delay Measurement Framework

We constructed our intrinsic light curve:

- Given Δ and θ, we have $\{\boldsymbol{t}^{\boldsymbol{\Delta}}, \boldsymbol{z}, \boldsymbol{\delta}^{\boldsymbol{z}}\}$ for the intrinsic light curve.
- Assumption: $\boldsymbol{z}$ = discrete realization of unobserved continuous process $z(t)$.

Let's model the stochastic variability in the intrinsic lightcurve:

- Parametric model, with parameter vector Ω, i.e. define $p(\boldsymbol{z}|\Delta, \boldsymbol{\theta}, \Omega)$
- Tak et al. [2017]: CARMA(1,0) (DRW) process.
- We generalize this to CARMA(p, q) processes.

Likelihood function of model parameters $(\Delta, \boldsymbol{\theta}, \Omega)$ given observed data

$$\begin{aligned} L(\Delta, \boldsymbol{\theta}, \Omega) &= p(\boldsymbol{x}, \boldsymbol{y}|\Delta, \boldsymbol{\theta}, \Omega) \\ &= p(\boldsymbol{z}|\Delta, \boldsymbol{\theta}, \Omega) \end{aligned} \quad (6)$$

DRW and CARMA(p, q) processes

- A **Damped Random Walk (DRW)** process with mean μ is the solution to the following stochastic differential equation:

$$\boxed{dX(t) = -\frac{1}{\omega}(X(t) - \mu)dt + \epsilon(t)} \tag{7}$$

  - $\epsilon(t) \sim N(0, \sigma^2)$ is a white noise process.
  - ω = timescale for the process to revert to its long-term mean
  - DRW = CARMA(1,0)

- A **CARMA(p, q)** process is the solution to the following stochastic differential equation:

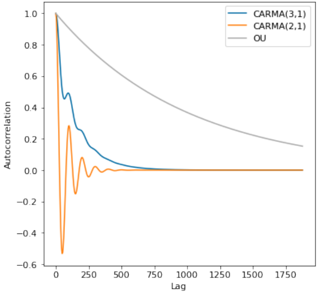
$$\boxed{\frac{d^p y(t)}{dt^p} + \alpha_{p-1}\frac{d^{p-1}y(t)}{dt^{p-1}} + \cdots + \alpha_0 y(t) = \beta_q \frac{d^q \epsilon(t)}{dt^q} + \cdots + \epsilon(t)} \tag{8}$$

  - $\boldsymbol{\alpha} = \{\alpha_0, \ldots, \alpha_{p-1}\}$ = auto-regressive parameters
  - $\boldsymbol{\beta} = \{\beta_0, \ldots, \beta_{q-1}\}$ = moving-average coefficients

DRW and CARMA(p, q) processes

In time-space (Auto-Correlation Function):

- DRW auto-correlations: single exponentially decaying auto-correlation function
- CARMA auto-correlations: weighted sum of exponentially decaying auto-correlation functions and exponentially damped sinusoidal functions



In frequency-space (Power Spectral Density)

- CARMA: multiple breaks and frequencies (QPOs) in the PSD

DRW and CARMA: Auto-Covariance Functions

ACF of a DRW process: $(t_i > t_j)$

$$R(t_i - t_j) = \frac{\sigma^2 \omega}{2} e^{-\frac{(t_i - t_j)}{\omega}} \tag{9}$$

- exponentially decaying with e-folding timescale ω

ACF of a CARMA(p, q) process: $(t_i > t_j)$

$$R(t_i - t_j) = \sigma^2 \sum_{k=1}^{p} \frac{\left[\sum_{l=0}^{q} \beta_l r_k^l\right] \left[\sum_{l=0}^{q} \beta_l (-r_k)^l\right] \exp(r_k \tau)}{-2\mathrm{Re}(r_k) \prod_{l=1, l \neq k} (r_l - r_k)(r_l^* + r_k)}. \tag{10}$$

- r_k = roots of auto-regressive polynomial $A(z) = \sum_{k=0}^{p} \alpha_k z^k$
- ACF of CARMA is a weighted sum of:
  - exponentially decaying components (when r_k is real)
  - exponentially damped sinusoids (when r_k is complex)
- Enforce $Re(r_k) < 0$ for stationarity

DRW and CARMA: Power Spectrum Density

PSD of a DRW process:

$$P(f) = \sigma^2 \frac{1}{(\frac{1}{\omega})^2 + (2\pi f)^2} \tag{11}$$

- Lorentzian centered at 0, with a break frequency at $1/2\pi\omega^2$

PSD of a CARMA process:

$$P(f) = \sigma^2 \frac{|\sum_{j=0}^{q} \beta_j (2\pi i f)^j|^2}{|\sum_{k=0}^{p} \alpha_k (2\pi i f)^k|^2} \tag{12}$$

- Weighted sum of Lorentzian functions
- Lorentzian centered at 0 $\rightarrow$ break frequency (when r_k is real)
- Lorentzian centered away from 0 $\rightarrow$ Quasi-Periodic Oscillation (QPO, when r_k is complex).

Characteristics such as multiple break-like features and QPOs have been observed in AGN optical data Kelly et al. [2014], Ryan et al. [2019]

Likelihood

CARMA and DRW process are Gaussian $\rightarrow p(z|\Delta, \theta, \Omega)$ **is Gaussian**

- Typically, the computation of the likelihood of an n-point realization of a Gaussian process requires the inversion of an $n \times n$ **covariance matrix** $\rightarrow$ **scales** $O(n^3)$**.**

$$\begin{aligned} L(\Delta, \boldsymbol{\theta}, \Omega) &= p(\boldsymbol{z}|\Delta, \boldsymbol{\theta}, \Omega) \\ &= \prod_{i=1}^{2n} p(z_i|\boldsymbol{z}_{<i}, \Delta, \boldsymbol{\theta}, \Omega) \\ &\propto \prod_{i=1}^{2n} \frac{1}{\mathrm{Var}(z_i|\boldsymbol{z}_{<i}, \Delta, \boldsymbol{\theta}, \Omega)} \quad \times \exp\left(-\frac{1}{2}\frac{(z_i - \mathrm{E}(z_i|\boldsymbol{z}_{<i}, \Delta, \boldsymbol{\theta}, \Omega))^2}{\mathrm{Var}(z_i|\boldsymbol{z}_{<i}, \Delta, \boldsymbol{\theta}, \Omega)}\right) \end{aligned}$$

- **But CARMA processes are special!**
- They admit a linear state-space representation that allows to compute the likelihood in linear time $O(n)$!

Computing the likelihood: state-space representation

Linear state-space representation of a CARMA(p,q) process denoted by $z(t)$ is:

$$\begin{cases} z(t) = \mathbf{b}\mathbf{x}(t) + \delta(t), \\ d\mathbf{x}(t) = A\mathbf{x}(t)dt + \mathbf{e}dW(t) \end{cases} \tag{13}$$

- Latent state process $\boldsymbol{x}(t)$ governs the underlying dynamics of the system
- Observation equation gives relationship between latent process $\boldsymbol{x}$ and observed process $z(t)$.
- $\delta(t)$ measurement error process.

Kalman Filter algorithm efficiently computes $\{\mathbf{E}(z_i|\boldsymbol{z}_{<i}, \Delta, \boldsymbol{\theta}, \Omega)\}_{i=1}^{n}$ and $\{\mathbf{Var}(z_i|\boldsymbol{z}_{<i}, \Delta, \boldsymbol{\theta}, \Omega)\}_{i=1}^{n}$ with linear complexity $O(n)$.

- CARMA processes modelling is scalable to large datasets

Bayesian Inference

We operate under the Bayesian paradigm:

- Quantify the uncertainty in model parameters via their (joint) posterior distribution
- Bayes' Theorem:

$$\boxed{p(\Delta, \boldsymbol{\beta}, \Omega | \boldsymbol{D}) = \frac{L(\Delta, \boldsymbol{\beta}, \Omega) p(\Delta, \boldsymbol{\beta}, \Omega)}{p(\boldsymbol{D})}} \tag{14}$$

- $p(\Delta, \boldsymbol{\beta}, \Omega | \boldsymbol{D}) \rightarrow$ posterior distribution
- $p(\boldsymbol{D}) \equiv \mathcal{Z} \rightarrow$ marginal distribution of the data (can be used for Bayesian model comparison).
- $p(\Delta, \boldsymbol{\beta}, \Omega) \rightarrow$ prior distribution (we choose uniform priors)

Prior distributions:

- $\Delta \sim [t_1 - t_n, t_n - t_1]$, $\mu \sim [-30, 30]$, $\theta \sim [-M, M]$ with M large
- α, β, σ sampled on log-scale, [-15,15].

Posterior sampling

We produce a sample of the posterior distribution using the MultiNest implementation of Nested Sampling (NS).

- Standard MCMC have trouble sampling from multi-modal posteriors
- MultiNest is designed to sample from multi-modal posterior posteriors

Most time delay estimation techniques are highly sensitive and require the input of an initial guess for Δ

- Need auxiliary method or prior knowledge to find plausible initial value for Δ
- Tak et al. [2017] compute expensive profile likelihood to find good initial guess.

MultiNest does not require the input of an initial value

- You just need to specify the boundaries of the parameter space!
- Method is blind search and standalone

Bayesian Model Selection

Model selection problem:

- what is the (p, q, m) triplet that best fits the data?
- CARMA(p, q) models are non-nested.

MultiNest evaluates the Bayesian evidence $\mathcal{Z}$:

$$\mathcal{Z} = \int_{\Theta} p(\boldsymbol{D}|\theta)p(\theta)d\theta \tag{15}$$

- $\mathcal{Z}$ is a measure of the "goodness of fit" of a model to the data $\boldsymbol{D}$.
- **Incorporates Occam's razor:** more complex models, i.e. models defined on higher-dimensional parameter spaces, are penalized if they do not sufficiently improve the fit to the data.
- No extra computational cost: $\mathcal{Z}$ is computed jointly to the posterior sampling

Tackling multi-modality with MultiNest

Two multi-modality issues:

- CARMA(p, q) parameters $\boldsymbol{\alpha}, \boldsymbol{\beta}, \sigma$
- Time Delay parameter Δ

Multi-modal posterior distributions are difficult to sample from. But MultiNest can help!

- MultiNest identifies modes in the posterior distribution
- Partitions the parameter space into regions Θ_i on which the separated modes are supported
- Evaluates the local-evidence $\mathcal{Z}_i$ of each mode i, defined as:

$$\mathcal{Z}_i = \int_{\Theta_i} p(\boldsymbol{D}|\theta)p(\theta)d\theta. \quad (16)$$

- Compute the relative probability p_i of mode i:

$$p_i \equiv \int_{\Theta_i} p(\theta|\boldsymbol{D})d\theta = \frac{1}{\mathcal{Z}} \int_{\Theta_i} p(\boldsymbol{D}|\theta)p(\theta)d\theta = \frac{\mathcal{Z}_i}{\mathcal{Z}}.$$

Time Delay Challenge dataset

Generated under the DRW model

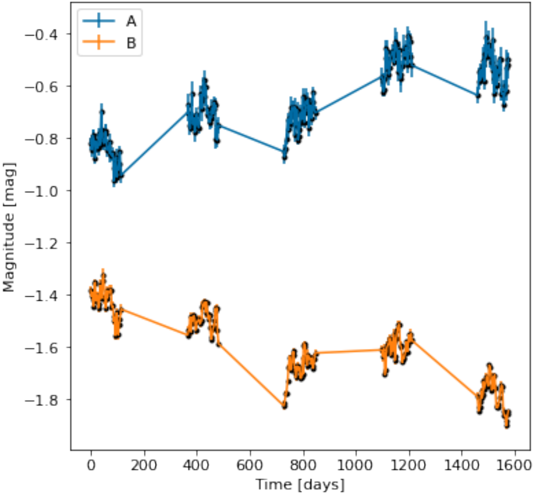


Figure: Doubly-lensed simulated quasar dataset from Time Delay Challenge (TDC)

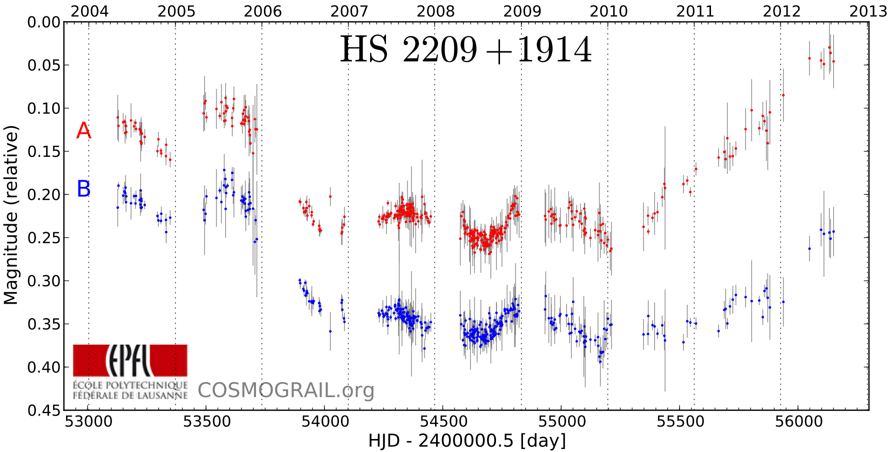
Time Delay Challenge (TDC) results

- We fit TD-CARMA(p, q, m) with $p = \{2, 3, 4\}$, $q = \{0, 1, 2, 3\}$ and $m = \{1, 2, 3\}$.

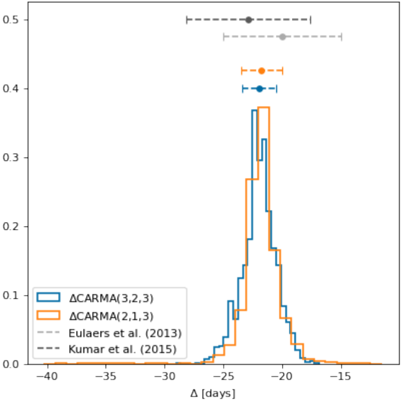
Model	$\hat{\Delta}$	SD($\hat{\Delta}$)	ln($\mathcal{Z}$)
Truth	*5.86*		
TD-DRW(3) [Tak et al., 2017]	6.33	0.28	
TD-DRW(3) (This work)	6.343	0.267	692.32
TD-CARMA(4, 3, 2)	6.296	0.255	699.11
TD-CARMA(4, 2, 2)	6.282	0.252	698.98
TD-CARMA(4, 1, 2)	6.287	0.259	698.49
TD-CARMA(2, 1, 2)	6.254	0.230	698.44
TD-CARMA(2, 0, 2)	6.269	0.242	698.39

Table: Posterior mean and standard deviation for Δ under the five models with highest Bayesian log-evidence, comparing to true value and estimates from DRW methods, as reported by Tak et al. [2017] and computed by our own code.

Dataset Application - HS2209



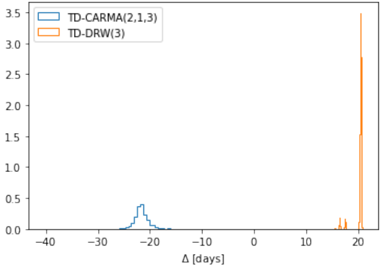
HS2209: Results



Technique	Reference	$\hat{\Delta}$	$SD(\hat{\Delta})$
Combined estimate (COSMOGRAIL)	Eulaers et al. [2013]	-20.0	5
Difference-smoothing (modified)	Kumar et al. [2015]	-22.9	5.3
$\Delta CARMA(3,2,3)$	This work	-21.96	1.448
$\Delta CARMA(2,1,3)$	This work	-21.74	1.423

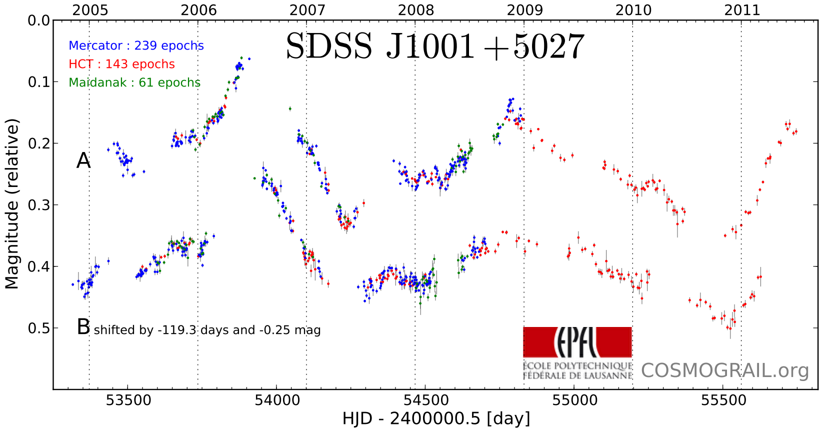
DRW process on HS2209 data

- Inconclusive results using *timedelay* package from Tak et al. [2017].
- With our code, MultiNest finds 10 modes in the posterior distribution of Δ. Modes for Δ include $[-14.30, -11.72, 16.44, 17.50, 20.38, 43, 64]$ days.



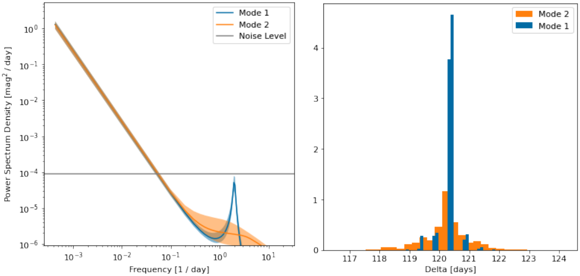
Model	$\hat{\Delta}$	SD($\hat{\Delta}$)	ln($\mathcal{Z}$)	p
TD-CARMA(3, 2, 3)	-21.96	1.448	2760.24	0.601
TD-CARMA(4, 2, 3)	-21.95	1.403	2759.83	0.399
TD-CARMA(2, 1, 3)	-21.74	1.423	2752.52	4.4×10^{-4}
TD-DRW(3)	20.23	0.934	2536.03	4.23×10^{-98}

SDSS J1001+5027 doubly lensed quasar



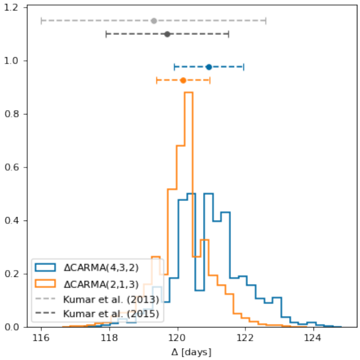
J1001: Multi-modality of CARMA parameters

- Two modes in the posterior distribution of CARMA parameters are identified by MultiNest.
- One mode corresponds to a frequency in the PSD ($f = 2$), but this frequency falls below the **measurement noise level** (so we are discarding modes/models that feature the frequency)



- Detection of frequency can dramatically reduce uncertainty in time delay Δ ($SD(\hat{\Delta}) = 0.686$ without freq, 0.224 with) $\rightarrow$ **but only if we believe the frequency exists!**

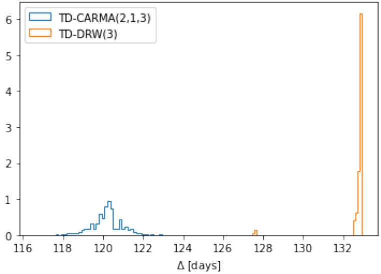
J1001: Results



Technique	Reference	$\hat{\Delta}$	$SD(\hat{\Delta})$
Combined estimate (COSMOGRAIL)	Kumar et al. [2013]	119.1	3.3
Gaussian Processes	Hojjati et al. [2013]	117.8	3.2
Difference-smoothing (modified)	Kumar et al. [2015]	119.7	1.8
ΔCARMA$(2,1,3)$	This work	120.18	0.749
ΔCARMA$(4,3,2)$	This work	120.93	1.015

DRW process on J1001 data

- Inconclusive results using *timedelay* package from Tak et al. [2017].
- With our code, MultiNest finds 20 modes in the posterior distribution of Δ. Modes for Δ include $[122.8, 127.6, 130.5, 132.8]$ days.



Model	$\hat{\Delta}$	SD($\hat{\Delta}$)	ln($\mathcal{Z}$)	p
TD-CARMA(4, 3, 2)	120.93	1.015	2761.25	0.416
TD-CARMA(2, 1, 3)	120.18	0.749	2744.05	4.1×10^{-8}
TD-OU(3)	132.71	0.750	1803.24	0.0

Future Research

Improvements of the method:

- Speed up likelihood computation using *celerite* model [Foreman-Mackey et al., 2017] (same complexity).
- Multi-band light curves?

Applications of the method:

- H_0 estimation
- Time delays arising in reverberation mapping

THANK YOU!

Hyungsuk Tak, Kaisey Mandel, David A Van Dyk, Vinay L Kashyap, Xiao-Li Meng, Aneta Siemiginowska, et al. Bayesian estimates of astronomical time delays between gravitationally lensed stochastic light curves. *The Annals of Applied Statistics*, 11(3):1309–1348, 2017.

Sjur Refsdal. On the possibility of determining hubble's parameter and the masses of galaxies from the gravitational lens effect. *Monthly Notices of the Royal Astronomical Society*, 128(4):307–310, 1964.

Kai Liao, Tommaso Treu, Phil Marshall, Christopher D Fassnacht, Nick Rumbaugh, Gregory Dobler, Amir Aghamousa, Vivien Bonvin, Frederic Courbin, Alireza Hojjati, et al. Strong lens time delay challenge. ii. results of tdc1. *The Astrophysical Journal*, 800(1):11, 2015.

Brandon C Kelly, Andrew C Becker, Malgosia Sobolewska, Aneta Siemiginowska, and Phil Uttley. Flexible and scalable methods for quantifying stochastic variability in the era of massive time-domain astronomical data sets. *The Astrophysical Journal*, 788(1):33, 2014.

Jackeline Moreno, Michael S Vogeley, Gordon T Richards, and Weixiang Yu. Stochastic modeling handbook for optical agn variability. *Publications of the Astronomical Society of the Pacific*, 131(1000):063001, 2019.

James L Ryan, Aneta Siemiginowska, MA Sobolewska, and Jonathan Grindlay. Characteristic variability timescales in the gamma-ray power spectra of blazars. *The Astrophysical Journal*, 885(1):12, 2019.

Eva Eulaers, Malte Tewes, Pierre Magain, Frédéric Courbin, Ildar Asfandiyarov, Sh Ehgamberdiev, S Rathna Kumar, CS Stalin, TP Prabhu, George Meylan, et al. Cosmograil: the cosmological monitoring of gravitational lenses-xii. time delays of the doubly lensed quasars sdss j1206+ 4332 and hs 2209+ 1914. *Astronomy & Astrophysics*, 553:A121, 2013.

S Rathna Kumar, CS Stalin, and TP Prabhu. H0 from ten well-measured time delay lenses. *Astronomy & Astrophysics*, 580:A38, 2015.

S Rathna Kumar, M Tewes, CS Stalin, F Courbin, I Asfandiyarov, G Meylan, Eva Eulaers, TP Prabhu, Pierre Magain, Hans Van Winckel, et al. Cosmograil: the cosmological monitoring of gravitational lenses-xiv. time delay of the doubly lensed quasar sdss j1001+ 5027. *Astronomy & Astrophysics*, 557:A44, 2013.

Alireza Hojjati, Alex G Kim, and Eric V Linder. Robust strong lensing time delay estimation. *Physical Review D*, 87(12):123512, 2013.

Daniel Foreman-Mackey, Eric Agol, Sivaram Ambikasaran, and Ruth Angus. Fast and scalable gaussian process modeling with applications to astronomical time series. *The Astronomical Journal*, 154(6):220, 2017.