

A kernel-based approach to estimating phase shifts between irregularly sampled time series: an application to gravitational lenses

Juan C. Cuevas-Tello^{1,3}, Peter Tiño¹, and Somak Raychaudhury²

¹ School of Computer Science, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom, {J.C.Cuevas, P.Tino}@cs.bham.ac.uk

² School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom, somak@star.sr.bham.ac.uk

³ Engineering Faculty, Autonomous University of San Luis Potosí, México

Abstract. Given two scaled, phase shifted and irregularly sampled noisy realisations of the same process, we attempt to recover the phase shift in this contribution. We suggest a kernel-based method that directly models the underlying process via a linear combination of Gaussian kernels. We apply our method to estimate the phase shift between temporal variations, in the brightness of multiple images of the same distant gravitationally lensed quasar, from irregular but simultaneous observations of all images. In a set of controlled experiments, our method outperforms other state-of-art statistical methods used in astrophysics, in particular in the presence of realistic gaps and Gaussian noise in the data. We apply the method to actual observations (at several optical frequencies) of the doubly imaged quasar Q0957+561. Our estimates at various frequencies are more consistent than those of the currently used methods.

Keywords: Kernel methods, time-series, regression

1 Introduction

According to the General theory of Relativity, a ray of light (or any other form of electromagnetic radiation, e.g. radio or x-rays) travels along a geodesic, which could be locally curved due to the gravitational effect of clumps of matter like stars or galaxies. This is known as Gravitational lensing [1] and gives rise to interesting cosmic illusions like magnified and seriously distorted images of distant sources, sometimes splitting into multiple images (e.g. Fig. 1), caused by intervening matter along the line of sight. Since the distortion of the images depends on the distribution of matter in the lensing object, this is the most direct method of measuring matter (which is often dark) in the Universe [2].

The quasar Q0957+561, an ultra-bright galaxy with a super massive central black hole (see Fig. 1), was the first lensed source to be discovered and it is the most studied so far. The source is 3.2×10^{10} light-years away from us, being lensed by a galaxy (visible in Fig. 1), along the line of sight, only 0.6×10^{10} light-years away. The effect of the lens is to create two distinct images of the

same source. The brightness of quasars varies on the time scales of days- and this variation shows up at different times in the two images since the path of light travel is different for them. Since such a *time delay* (phase shift) can provide a rare direct measure of the distances involved, this quantity is of great importance in astronomy, and thus it is not surprising that many attempts have been made to estimate it, e.g. see [3–6].

The observations can be made by both radio and optical astronomers, since theory predicts that the time delay is independent of the frequency of observation. For our purposes, the data are available as two unevenly sampled time series of fluxes (or logarithm thereof) of the two images. The observations are made at irregular intervals due to weather conditions, equipment availability, object visibility, among other practical considerations.

Elsewhere, we have empirically shown, using artificial irregularly sampled time series with noise and gaps (typical of radio observations), that a kernel-based approach to measure the time delay between two such time series outperforms typical statistical methods used by astronomers [7]. In this contribution, we extend, improve and test this approach to analyse actual optical observations (and artificial data representing such data), which shows high variability compared with radio observations. We compare results with the dispersion spectra method [3, 4], which is the most reliable of methods used by astronomers, and thus very widely used [8].

The remainder of this paper is organised as follows: §2 describes the optical data, and §3 the methods. Results of our method are compared to those from the dispersion spectra method in §4, and in §5 we show results from our analysis of optical-like artificial data, followed by comments and conclusions.

2 Astronomical Observations

In this work, we analyse the brightness of the two images of quasar Q0957+561 (Fig. 1) as a function of time, to find the phase shift between the time series. The data sets analysed here are summarised in Table 1. Optical astronomers measure the brightness of a source using imaging devices, with filters to restrict the range of wavelength/frequency of light observed. The flux f of light from a source is expressed in logarithmic units known as magnitudes (mag), defined as $\text{mag} = -2.5 \log_{10} f + \text{constant}$. The errors on mag mainly measurement errors, assumed to be zero-mean Gaussian. The green (g) and red (r) bands represent measurements obtained with filters in the wavelength range 400–550 nm and 550–700 nm, respectively. We use the data sets DS1 and DS2 [5], obtained through a monitoring program at the Apache Point Observatory, New Mexico, USA, and DS3, from images taken at Fred Lawrence Whipple Observatory, Mt. Hopkins, Arizona, USA. [6]. The results of this analysis are presented in §4.

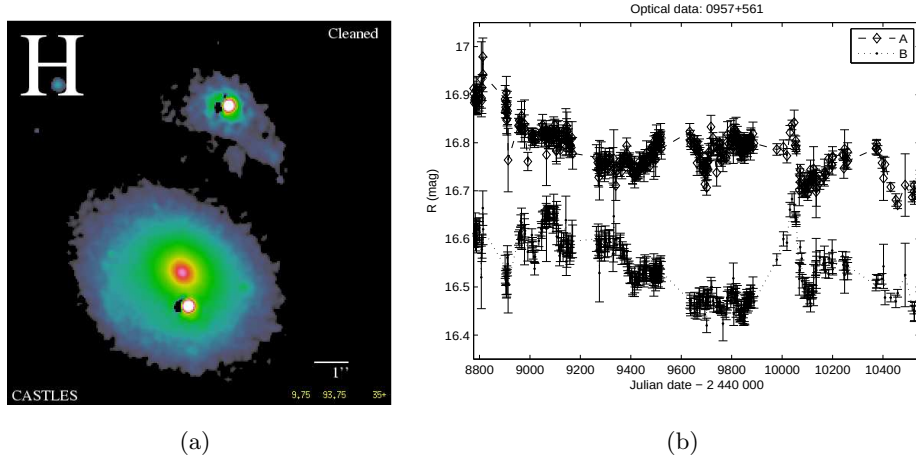


Fig. 1. Quasar Q0957+561. **(a)** Image taken by the Hubble Space telescope (<http://www.cfa.harvard.edu/castles>). The two point images are of the same distant quasar, 32 billion light-years away, multiply-imaged due to the gravitational effect of the “lensing” galaxy, seen as the extended object, which is along the line of sight, 6 billion light-years away from us. **(b)** The two time series represent the brightness of the two images (in logarithmic units (mag), such that brighter means lower values; see text) as a function time (the abscissa is measured in *days*). Image A is shifted up by 0.2 mag for visualisation purposes. This is data set DS3 with measurement error bars (std. deviations), see §2 and Table 1 for details.

3 Methods for time delay estimation

We model a pair of time series, obtained by monitoring the brightness (in mag units) of image A and image B, as follows

$$\begin{aligned} x_A(t_i) &= h_A(t_i) + \varepsilon_A(t_i) \\ x_B(t_i) &= h_B(t_i) + M + \varepsilon_B(t_i), \end{aligned} \quad (1)$$

where M is the offset between the two images, and $t_i, i = 1, 2, \dots, n$ are discrete observation times. The observation errors $\varepsilon_A(t_i)$ and $\varepsilon_B(t_i)$ are modelled as zero-mean Normal distributions $N(0, \sigma_A(t_i))$ and $N(0, \sigma_B(t_i))$, respectively; $\sigma_A(t_i)$ and $\sigma_B(t_i)$ are given. Now,

$$h_A(t_i) = \sum_{j=1}^N \alpha_j K(t_j, t_i) \quad (2)$$

is the “underlying” light curve that underpins image A, whereas

$$h_B(t_i) = \sum_{j=1}^N \alpha_j K(t_j + \Delta, t_i) \quad (3)$$

Table 1. 0957+561 optical data sets analysed here

id	band	# Samples	Date	Reference
DS1	<i>g</i>	97	2/12/94 to 6/7/96	[5]
DS2	<i>r</i>	100	2/12/94 to 6/7/96	[5]
DS3	<i>r</i>	422	2/6/92 to 8/4/97	[6]

is a time-delayed (by Δ) version of $h_A(t_i)$ underpinning image B. The Gaussian kernels $K(\cdot, \cdot)$ are centred at either t_j , $j = 1, 2, \dots, N$ (function f_A), or $t_j + \Delta$, $j = 1, 2, \dots, N$ (function f_B) [9, 10]. We use widths $\omega_{c_j} > 0$ determining the ‘degree of smoothness’ of the models h_A and h_B . The widths $\omega_j \equiv \omega_{c_j}$ are determined through the k nearest neighbours of t_j as follows:

$$\omega_j = \sum_{d=1}^k (t_j - t_{j-d}) + (t_{j+d} - t_j) = \sum_{d=1}^k (t_{j+d} - t_{j-d}).$$

The value of parameter k can be estimated via cross validation.

The weights α in (2-3) are given by

$$\mathbf{K}\alpha = \mathbf{x}, \quad (4)$$

where $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_N)^T$,

$$\mathbf{K} = \begin{bmatrix} K_A(\cdot, \cdot) \\ K_B(\cdot, \cdot) \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_A(\cdot)/\sigma_A(\cdot) \\ x_B(\cdot)/\sigma_B(\cdot) \end{bmatrix}, \quad (5)$$

and the kernels $K_A(\cdot, \cdot)$, $K_B(\cdot, \cdot)$ have the form [7]:

$$K_A(t_j, t_i) = \frac{K(t_j, t_i)}{\sigma_A(t_i)}, \quad K_B(t_j, t_i) = \frac{M + K(t_j + \Delta, t_i)}{\sigma_B(t_i)}. \quad (6)$$

Our aim is to estimate the time delay Δ between the temporal light curves corresponding to images A and B. Given the observed data and a suggested delay Δ ($[\Delta_{min}, \Delta_{max}]$), free parameters of the model (1-3) are determined within the maximum likelihood framework. Since the model is linear in parameters, we regularise $\mathbf{K}$ in the model fitting via singular value decomposition (SVD) [11, 12].

We use model formulation (1-3), because **(1)** linearity in parameters enables us to use tools of linear algebra in parameter fitting and regularisation, **(2)** Gaussian kernel formulation using variable kernel widths is natural in cases of irregularly sampled data, **(3)** parameter sharing in (2) and (3) provides a transparent tool for coupling the two observed images.

To measure the time delay between time series, astrophysicists often use the Dispersion, which is a weighted sum of squared differences between $x_A(t_i)$ and

$x_B(t_i)$. We use the D_1^2 [3] and $D_{4,2}^2$ [4] methods. The latter has a free parameter, *decorrelation length* δ , that signifies the maximum distance between observations we are willing to consider when calculating the correlations. The estimated time delay, Δ , is found by minimising D^2 over a range of time delay trials $[\Delta_{min}, \Delta_{max}]$.

4 Analysis of optical monitoring of gravitational lens Q0957+561

We apply both the Dispersion method, which astrophysicists commonly use, and our method on the observational data sets, summarised in Table 1, consisting of measures of the brightness of the two images at irregular intervals.

For the time delay, we use bounds of $\Delta_{min} = 400$ and $\Delta_{max} = 450$ days given that our prior knowledge (from other analyses) is that the best delay is around 420 days [5, 6]. So, we evaluate D_1^2 and $D_{4,2}^2$ in this range with increments of one day. The results are in Table 2. The decorrelation length δ in Table 2 is the same adopted by [5] and [6].

The confidence intervals were estimated through 500 Monte Carlo simulations over the observation noise processes by fixing the parameters M and δ to the best values, as in Table 2. The results are in Table 3, where μ_Δ is the mean of the time delay and the confidence intervals are given by the standard deviations (σ_Δ).

Table 2. Results on observed data

Data set	Dispersion spectra		Kernel-based approach
	$D_1^2: \Delta (M)$	$D_{4,2}^2: \Delta (M; \delta)$	$\Delta (k)$
DS1	417 (0.119)	420 (0.109;7)	420 (3)
DS2	429 (0.210)	446 (0.210;7)	420 (3)
DS3	425 (0.077)	424 (0.077;4)	430 (6)

Quantities in days

When applying our model, we have fixed M to 0.117, 0.21 and 0.076 for DS1, DS2 and DS3 respectively [5, 6]. Singular values of $\mathbf{K}$ less than a threshold $\lambda = 0.001$ are set to zero to avoid ill-conditioning [11, 12] and the smoothing parameter k was chosen through five-fold cross validation (CV) [7]. The results are in Table 2. Again, confidence intervals are estimated through 500 Monte Carlo simulations fixing M , k and λ to their optimal values (see Table 3).

5 Artificial data

Since the true time delay on the quasar Q0957+561 is unknown, the best way to compare the performance of methods is through a set of controlled experiments

Table 3. Confidence intervals: 500 Monte Carlo simulations

Data set	Dispersion spectra		Kernel-based approach	
	$D_1^2 : \mu_\Delta \pm \sigma_\Delta$	$D_{4,2}^2 : \mu_\Delta \pm \sigma_\Delta$	$\mu_\Delta \pm \sigma_\Delta$	k
DS1	416.7 $\pm$ 0.9	419.9 $\pm$ 1.3	419.5 $\pm$ 0.8	3
DS2	421.6 $\pm$ 2.8	443.5 $\pm$ 8.2	421.3 $\pm$ 3.6	3
DS3	426.7 $\pm$ 2.3	438.5 $\pm$ 12.7	432.2 $\pm$ 5.3	6

μ_Δ and σ_Δ are given in days

where the true time delay is known. We use optical-like artificial data to compare our approach with the commonly used dispersion spectra method. In [7], we used radio-like artificial data with an imposed time delay of 500 days over an observational season of 13.6 years.

Here, the artificial data is generated as in [7], but with an observational season of 1.3 years, 50 irregular samples, a true time delay of 5 days, an offset $M = 0.1$ (considered fixed and known for both methods in §3), three levels of noise of 0.03%, 0.106% and 0.466% of mag (minimum, average and maximum of DS3, respectively), and "observational" gap size of zero to five continuously missing samples per block (five blocks randomly located). We use ten different underlying functions⁴, 100 realisations per level of noise and ten realisations per gap size. This gives us an amount of 153,510 data sets under analysis. So, these data sets simulate optical data with low time delay and low offset with high precision [6].

The results are in Fig. 2, mean and standard deviation of time delay estimates are calculated for each underlying function. Then, the mean and standard deviation across all artificial data sets, μ_Δ and σ_Δ , respectively, are calculated and depicted in Fig. 2. We stress that to make our comparison fair, each method was subjected to the same collection of artificial data sets. In all cases the time delay under analysis is from 0 to 10 days; with increments of 0.1 days. The parameter λ is fixed as above.

6 Comments and conclusions

On Monte Carlo simulations, the set DS1 leads to the minimum standard deviation for both dispersion spectra methods, as well as for our kernel-based approach; see Table 3. With our methods, we get consistent results for DS1 and DS2 in Table 2, because they have almost the same sampling. On the other hand, Kundic et al. did not find such a concord with the four methods studied in [5]. Rather, they adopted the time delay of 417 ± 3 days given by *Linear method* [5]. Therefore, the best time delay for DS1 and DS2 is 420 days rather than 417 days [5]. Nevertheless, nobody knows the true time delay for the quasar Q957+561 so far, and as more observations are gathered more time delay estimates appear.

⁴ plots available at <http://www.cs.bham.ac.uk/~jcc/artificial-optical/>

Therefore, in Fig. 2, we have a comparison of our approach against dispersion spectra on artificial data, where the true time delay is known ($\Delta = 5$ days). It appears that the D_1^2 method is less biased than the $D_{4,2}^2$ method. However, compared with $D_{4,2}^2$, the variance of D_1^2 estimates is higher. Compared with D_1^2 , our method has less bias and less variance, except for cases of 0% of noise and gap size less than 3, where we observe smaller bias but higher variance. Overall, compared with our method, D_1^2 and $D_{4,2}^2$ seem more vulnerable to observational gaps.

Based on the results in Tables 2 and 3, and in Fig. 2, we conclude that our method is more accurate than dispersion spectra (see caption to Fig. 2). In the future we also plan to investigate options for speeding up parameter estimation in our kernel-based approach.

References

1. Saha, P.: Gravitational Lensing. Encyclopedia of Astronomy and Astrophysics (2000)
2. Kochanek, C., Schechter, P.: The Hubble constant from gravitational lens time delays. Carnegie Observatories Astrophysics Series **2** (2004)
3. Pelt, J., Kayser, R., Refsdal, S., Schramm, T.: Time delay controversy on QSO 0957+561 not yet decided. Astronomy and Astrophysics **286**(1) (1994) 775–785
4. Pelt, J., Kayser, R., Refsdal, S., Schramm, T.: The light curve and the time delay of QSO 0957+561. Astronomy and Astrophysics **305**(1) (1996) 97–106
5. Kundic, T., Turner, E., Colley, W., Gott-III, J., Rhoads, J., Wang, Y., Bergeron, L., Gloria, K., Long, D., Malhorta, S., Wambsganss, J.: A robust determination of the time delay in 0957+561A,B and a measurement of the global value of Hubble’s constant. Astrophysical Journal **482**(1) (1997) 75–82
6. Ovaldsen, J., Teuber, J., Schild, R., Stabell, R.: New aperture photometry of QSO 0957+561; application to time delay and microlensing. Astronomy and Astrophysics **402**(3) (2003) 891–904
7. Cuevas-Tello, J., Tiño, P., Raychaudhury, S.: How accurate are the time delay estimates in gravitational lensing? Astronomy and Astrophysics (2006) Accepted, <http://arxiv.org/abs/astro-ph/0605042>.
8. Eigenbrod, A., Courbin, F., Vuissoz, C., Meylan, G., Saha, P., Dye, S.: COSMOGRAIL: The COSmological MONitoring of GRAvItational Lenses. I. How to sample the light curves of gravitationally lensed quasars to measure accurate time delays. Astronomy and Astrophysics **436** (2005) 25–35
9. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer (2001)
10. Shawe-Taylor, J., Cristianini, N.: Kernel Methods for Pattern Analysis. Cambridge University Press (2004)
11. Press, H., Flannery, B., Teukolsky, S., Vetterling, W.: Numerical Recipes. Cambridge University Press (1986)
12. Golub, G., Van-Loan, C.: Matrix Computations. second edn. The Johns Hopkins University Press (1989)

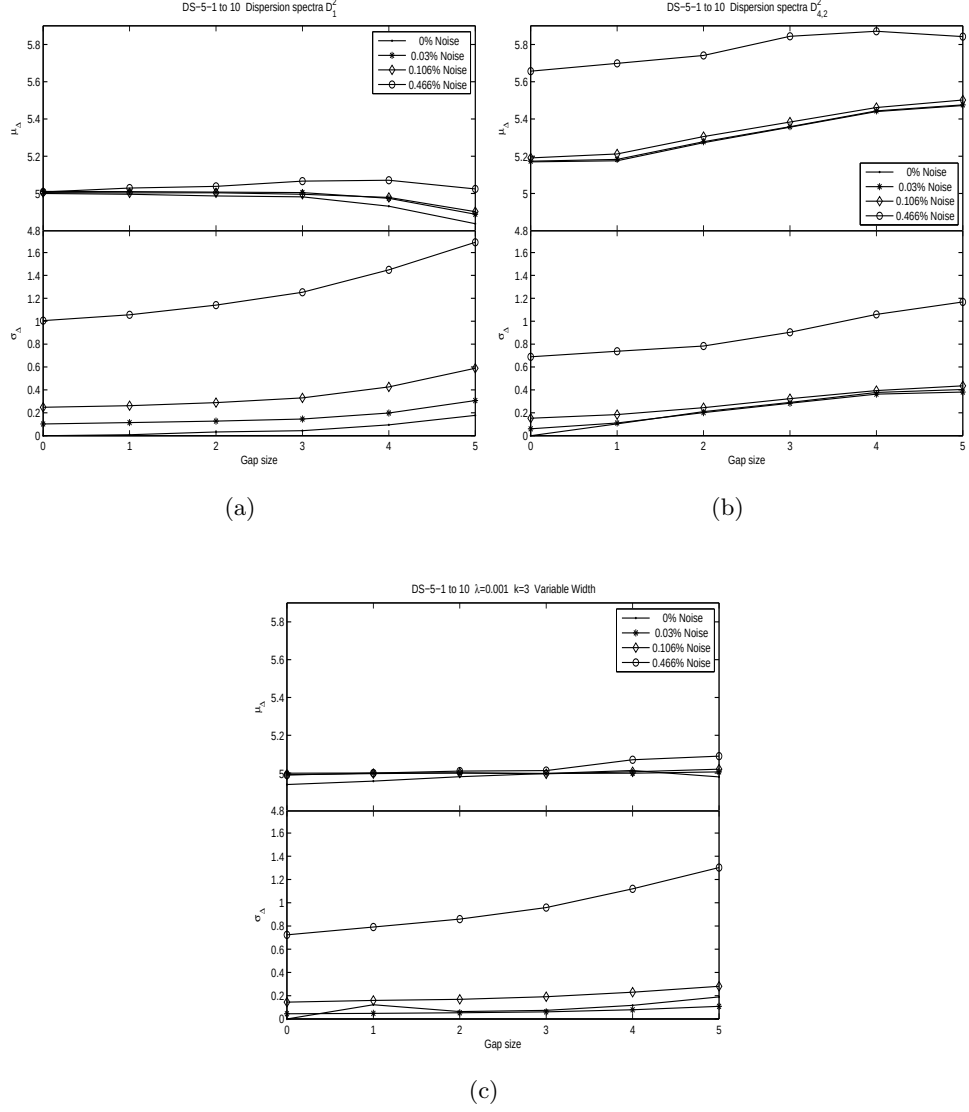


Fig. 2. Results on all artificial data, see §5 for details. **(a)** Dispersion spectra D_1^2 : values of μ_Δ range in $[4.83, 5.07]$, and for σ_Δ in $[0, 1.68]$. **(b)** Dispersion spectra $D_{4,2}^2$: decorrelation length δ was fixed to 5. The values of μ_Δ range in $[5.17, 5.87]$, and for σ_Δ in $[0, 1.16]$. **(c)** Kernel-based approach: parameter k was fixed to 3, and the regularisation parameter λ to 0.001. Values of μ_Δ range in $[4.94, 5.08]$, and for σ_Δ in $[0, 1.30]$.