PX4128 / PXT125 – Data Analysis

Mini Project: 40% of final grade
Deadline for submission: 2pm on Friday 15th December 2017

Note: Please read this whole assessment sheet and instructions several (many!) times to get a feel for what the project is asking of you. You might then want to make your own rough notes that outline how should proceed. The steps outlined below are there to help guide you through the process. **This is not a first year lab**: if you don't understand **why** you are being asked to do the step, you are probably not going to do it correctly. As 4th year students, you are now researchers! :)

Hunting a Black Hole in GS2000

In the close binary system GS2000, a late-type main-sequence star orbits around a compact object surrounded by an accretion disk. GS2000 is a "soft X-ray transient" discovered during an outburst of the disc when it became a bright X-ray source due to temporarily increased rate of accretion onto the compact object. The compact object is suspected to be a black hole, but it could also be a neutron star. Theory suggests that the maximum mass possible for a neutron star is about 3 solar masses.

The goal of this project is to use data analysis techniques learned in the lectures to estimate the mass of the compact object in the astronomical source GS2000, and hence to determine if it is likely to be a black hole. To measure the compact object mass, the strategy is to measure the period and orbital velocity of the companion star, and then apply Kepler's law. The orbital velocities can be measured via the red and blue doppler-shifting of the absorption lines in the companion star's spectrum, as it orbits around the compact object. You have two main problems: the star's spectra are very noisy (and contaminated with light from the accretion disc), and you don't exactly know what the spectrum of the star should look like.

The 10-m Keck telescope on Mauna Kea in Hawaii has been used to obtain 13 spectra of GS2000 at 13 different binary phases. The spectra include light both from the companion star and from the accretion disk. They were taken a few years after the outburst, when accretion disc had faded considerably, so that the companion star makes a substantial contribution to the spectrum.

Because GS2000 is quite faint, spectra are rather noisy even with with CCD spectrograph on the Keck telescope. Also, our spectra are unable to resolve the individual components of the system — instead of separate spectra for the compact object's accretion disc and companion star, we get a single merged spectrum at each observation. The Hα emission line from the (now) quiescent accretion disc around the compact object is readily visible, with a characteristic double-peaked velocity profile, but it is not easy to spot absorption lines from the companion star in the individual spectra. For this reason a cross-correlation analysis is needed to detect the radial velocity of the companion star at each of the 13 phases. A number of bright main sequence stars were observed to provide "template" spectra, in which the locations of absorption lines are easily visible.

The techniques required have been discussed in the lectures, although you may want to take another look at some of the course textbooks and do a little background reading to be sure you understand the statistics and physics of the problem. I assure you that, other than the jargon, the physics is mainly 1st year syllabus combined with some school trigonometry. The basic steps of the cross-correlation radial velocity analysis are out-lined below.

1. Stellar types (e.g. G4, M3, etc) encode the mass of the star. Spectral analysis of this system suggests that the companion star (the one you're trying to find in the data) is around

K5, but spectra are so noisy that this is very uncertain. However, this is probably a good starting point! Open up the K5 template file, and make a plot of the data. Also open up the GS2000 data and make a series of plots to see how the data vary with time. Detailed in formation on what is in the data files can be found in the Mini Project directory alongside the data files themselves. **Examine the data carefully before you start this project.**

2. For each of the 13 spectra, measure the radial velocity by performing a 'cross-correlation' analysis with your choice of template spectrum. Basically, the idea is to chose many 'trial' values for the orbital velocity, and see which one 'fits' the data the best, as measured by the chi-squared values. This is a 'brute force' approach! To accomplish this:

   (a) Remove the background continuum so that we just have a series of lines in each spectrum.You can do this by fitting a function to the overall shape of the spectra, and then subtract this from the data. You need to do this for the GS2000 spectrum, **and** from the template spectrum.

   (b) Doppler shift your continuum-subtracted template spectrum for the chosen trial velocity.

   (c) **Scale** the shifted and continuum-subtracted template spectrum to fit the continuum-subtracted GS2000 spectrum.

   (d) Repeat steps b) and c) for a number of different velocities, and plot the χ^2 of the fit as a function of the velocity shift.

   (e) Use the χ^2 minimum to estimate the radial velocity and its 1-sigma uncertainty.

3. Plot the measured velocities and 1-sigma uncertainties as a function of binary phase. Put these results also in a table in your report.

4. Fit a sinusoidal velocity curve of the form

$$V(\phi) = \gamma + K_x \sin(2\,\pi\,\phi) + K_y \cos(2\,\pi\,\phi)$$

   to the 13 velocity measurements. Show the best fit velocity curve on your plot with the data points and error bars.

5. From your results, estimate the radial velocity semi-amplitude,

$$K = \left(K_x{}^2 + K_y{}^2\right)^{1/2}$$

   and its 1-σ uncertainty. This is the orbital velocity of the companion star, projected by a factor sin (i) due to the inclination of the orbit.

6. The compact object (i.e. the blackhole / neutron star) mass function $f(M_X)$ is given by,

$$f(M_X) = \frac{M_X^3 \sin^3 i}{(M_X + M_C)^2} = \frac{M_X \sin^3 i}{(1+q)^2} = \frac{PK^3}{2\pi G}$$

   the first two expressions are in terms of the physical properties of the system — i.e.the compact object mass M_X, the inclination, i, of the orbital plane, and the mass ratio $q = M_C/M_X$, where M_C is the companion star mass — and the last expression holds the observable quantities P (period) and K. From the binary period P and your measured K velocity, estimate $f(M_X)$ in solar units. Give also the 1-sigma uncertainty.

7. Adopt a specific inclination $i = 90^{\circ}$, and a specific M_C based on the spectral type of the template star used. For these calculate the compact object mass M_X. Give the 1-σ uncertainty, assuming that i and M_C are known exactly.

8. In fact i and M_C are not known, but have some limited range. Use Monte-Carlo error propagation to sample the range of uncertainty in i, M_C, K, and P. Be careful to assume a distribution for i that corresponds to a random orbit orientation. For each Monte-Carlo trial, calculate the corresponding M_X. Make plots showing the resulting probability distribution for M_X, both as a probability density histogram, and as a cumulative probability function. Draw vertical lines on the plots to indicate the mean and median of the M_X distribution, and a 1-sigma confidence interval for M_X.

9. Give the probability that the compact object mass is higher than 3 solar masses, and hence the probability that it is a black hole rather than a neutron star.

Important Tip! Do not start coding straight away. Plan! Write out a sketch of what techniques you will need to apply, and what pieces of code you will need build/use. Then try to write pieces of code that do one simple thing (e.g. apply a doppler shift to a spectrum, or scale a function to fit data, etc). Once these work (i.e. you've tested them and understand the output!), then try to piece the functions together to to solve the problem. Be careful to inspect the data the carefully before and after going into your code.

All the data will be made available via the Learning Central portal. An email will be sent out once the data has been uploaded.

Grading

1) The piece of work will be marked out of 100.
2) 60 marks will be available for getting the basic analysis correct, and will be split up among the various stages outlined above. Be careful with these steps (in particular step 2!).
2) 25 marks will be available for more rigorous error analysis.
3) 15 marks will be available for the quality of the report, including clarity, ease of reading, and train of narrative.

Please use off-the-shelf random number generators and plotting routines. It is also OK to use off-the-shelf spline functions. However, please **write your own fitting algorithms**. Although the steps are outlined above, you need to say **why** each step is being done, to show that you understand the above procedure. Show all your working, coding, and produce a tidy, **succinct**, report. Full marks cannot be awarded if there is no code given in the report, and note that the code should be well commented. Please also be sure to cite any sources of information that you use in the report.

You must submit an **electronic copy** via the Turn-It-In post on Learning Central. Note that you **can only post one document to Turn-It-In**, so everything will need to be in a single PDF, including all your coding!