Technical Note

Investigation of the accuracy of cross-correlation to derive radial velocities from 4MOST spectra

Document No.: MST-TNO-XXX-YYYYY-GGGG-HHHH

Issue No.: 1.00

DocuShare Document URL: [Kommentare]

Prepared by 2019-01-11

Dominic Ford

Change Record

Issue	Date	Page, Section, Paragraph Affected	Reason, Remarks
1.00	2019-01-11	All	First version of document

List of Contributors

Name	First Name	Institute	Comments
Ford	Dominic	Lund Observatory	Lead Developer
Church	Ross	Lund Observatory	
Kordopatis	Georges	Observatoire de la Côte d'Azur, Nice	IWG7 lead
Lind	Karin	Max-Planck-Institut für Astronomie, Heidelberg	IWG7 lead
Feltzing	Sofia	Lund Observatory	Project Scientist
Traven	Gregor	Lund Observatory	

Table of contents

List of Figures

1 Scope

We present a study of the feasibility of using cross correlation to derive the radial velocities of stars from 4MOST observations.

We demonstrate that when mock observations are made of synthetic spectra which have had radial velocities applied to them, and these are cross-correlated with a library of templates, it is possible to recover the radial velocity to within the required precision stated in the IWG7 Management Plan [AD1].

We further demonstrate that when radial velocities are injected into the PEPSI observations of the Gaia Benchmark Stars, the same cross-correlation framework is again able to recover the radial velocity to within the required precision.

2 Applicable Documents (AD)

The following applicable documents (AD) of the exact issue shown form a part of this document to the extent described herein. In the event of conflict between the documents referenced herein and the contents of this document, the contents of this document are the superseding requirement.

AD ID	Document Title	Document Number	Issue	Date
[AD1]	IWG7 Management Plan	MST-PLA-PSC-20307-09237-0001	0.07	05.01.2017

3 Reference Documents (RD)

The following reference documents (RD) contain useful information relevant to the subject of the present document.

RD ID	Document Title	Document Number	Issue	Date
[RD1]	Ford, D. 2017a, The IWG7 4MOST Galactic Pipeline (4GP): Development Report (July 2017)			
[RD2]	Ford, D. 2017b, The IWG7 4MOST Galactic Pipeline (4GP): Development Report 2 (December 2017)			
[RD3]	Ford, D. 2018a, The IWG7 4MOST Galactic Pipeline (4GP): Development Report 3 (June 2018)			
[RD4]	Buder S. et al. 2018, MNRAS – http://adsabs.harvard.edu/doi/10.1093/mnras/sty1281			

4 Introduction

The IWG7 Management Plan [AD1] sets out a requirement that the 4MOST Galactic Pipeline (4GP) must be able to derive radial velocities from stellar spectra observed with 4MOST HRS with a precision of 1 km/s, and from LRS spectra with a precision of 2 km/s.

In this document, we implement a simple algorithm which estimates the radial velocities of stellar spectra by cross-correlating them with a small library of template spectra. We test its accuracy by injected radial velocities into the synthetic spectra and creating mock 4MOST observations of them using the 4FS exposure time calculator. The algorithm tested here is similar to the one used in the GALAH pipeline – specifically, the Galah Ultra-Enigmatic Spectroscopic Script (GUESS).

5 Method

The tests presented in this report were conducted within the framework of the 4MOST Galactic Pipeline (4GP). This is a collection of Python modules which provide a modular and extensible framework for testing the performance of algorithms in the reduction of stellar spectra. They have been described in more detail in previous technical notes (Ford et al. 2017a, 2017b, 2018a).

In particular, the cross-correlation code tested in this report is implemented by the `RvInstanceCrossCorrelation` class provided by the `fourgp_rv` package within 4GP.

The 4GP framework is available in two repositories on GitHub. The first repository contains Python modules which provide programmatic interfaces for manipulating libraries of spectra, including the `RvInstanceCrossCorrelation` class:

https://github.com/dcf21/4most-4gp

The second repository contains python scripts which provide command-line interfaces for performing a wide variety of tests and manipulations on spectra:

https://github.com/dcf21/4most-4gp-scripts

The scripts used to perform the tests described in this report are here:

https://github.com/dcf21/4most-4gp-scripts/tree/dev/src/scripts/rv_code_test

Step-by-step instructions for installing and running 4GP can be found in the Wiki pages on GitHub:

https://github.com/dcf21/4most-4gp/wiki
https://github.com/dcf21/4most-4gp-scripts/wiki

5.1 Generating template spectra

To define a small library of template spectra for use in the cross-correlation process, we chose a coarse rectangular grid of stellar parameters:

- $T_{\rm eff}$ between 4000 and 8000K, at 500K intervals
- log(g) between 1.5 and 4.5 dex, at 0.5 dex intervals
- [Fe/H] between -1.5 and 0.5 dex, at 1 dex intervals

We used Turbospectrum to create synthetic spectra for each of these 315 sets of stellar parameters, setting all elemental abundances to scaled-solar values. Turbospectrum is only able to synthesise spectra for stellar parameters which lie within the MARCS grid of model stellar atmospheres. Many of the sets of stellar parameters in the grid defined above are highly unphysical, and so model atmospheres are not available for them. In practice, this excludes over two-thirds of the points in this grid.

When these unphysical templates are excluded, a grid of 75 sets of stellar parameters remain.

Once Turbospectrum had generated high-resolution synthetic spectra for these parameters, we passed them through the 4FS exposure time calculator (ETC) to create mock 4MOST observations from them. This reduces the resolution of the spectra to match the 4MOST instrumental profile and injects a small amount of noise into them such that their SNR/pixel in the 6100 – 6600 Å region is 250.

5.1.1 Resampling onto a fixed logarithmic raster

The cross-correlation algorithm works by looking for a fixed offset, in pixels, between the test spectrum and the library of template spectra. This means that applying a radial velocity to a spectrum must have the effect of shifting it by a constant number of pixels. This means that the spectrum must be sampled at wavelengths with a fixed logarithmic stride. In other words, it must be possible to write the wavelength λ_i of pixel i as

$$\lambda_i = \lambda_0\, \alpha^i$$

for constant α. This is not true for the mock observations produced by 4FS, and so after passing each spectrum through 4FS, each of the three arms of the spectrum is subsequently resampled onto such a raster. The multiplicative step α is chosen to match the mean multiplicative step in the arm's original output from 4FS. The step is different for each of the three arms within 4MOST's high- and low-resolution spectrographs, and so the cross-correlation algorithm must be applied to each arm separately.

5.1.2 Edge effects

The cross-correlation technique is vulnerable to edge effects as it treats each spectrum as a repeating periodic signal. When the spectrum is shifted by some number of pixels, the pixels that fall off one end of the spectrum re-enter at the opposite end of the spectrum.

To minimise the effect of this, we multiply each template spectrum by a window function which tapers the edges to zero. Specifically, the 10% of pixels at either end of the spectrum are multiplied by a linear ramp function, which is zero for the end pixel, and reaches unity 10% of the way into the spectrum.

5.2 Generating test spectra

To create a physically plausible population of synthetic spectra which could be used to test the cross-correlation code, we took randomly selected objects from the GALAH test sample (Buder et al. 2018), as described in Ford et al. (2018a). We passed the stellar parameters and elemental abundances of these stars into Turbospectrum to generate a population of high-resolution synthetic spectra. We then applied random RVs to these synthetic spectra, with the following distribution:

- 90% were from a Gaussian with standard deviation 25 km/s
- 10% were from a uniform distribution between +/- 200 km/s

We passed the shifted spectra through the 4FS exposure time calculator to create mock observations of them. Following the procedure described above, we resampled each arm of the test spectra onto the same wavelength raster as the template spectra.

5.3 Cross-correlation

To determine the RV of each test spectrum, we compute the zero-mean normalised cross-correlation (ZNCC) function between each test spectrum and each of the 75 template spectra in turn. This is done for each arm of the spectra individually. For a discretely sampled spectrum, the ZNCC is defined as

$$(f \times g)[i] = \frac{\sum_j f[j]g[j+i]}{n}$$

where $f[i]$ and $g[i]$ are ith pixel in a renormalized version of the test spectrum and the template spectrum respectively, and n is the number of pixels being summed over. Each spectrum must be normalised such that the mean value of the pixels within it is zero, and their standard deviation is unity. This is achieved by calculating the mean and standard deviation of the pixel values within each spectrum, subtracting the mean from the value of each pixel, and then dividing it by the standard deviation.

With this normalisation, each pixel within the cross-correlation function equals the Pearson correlation coefficient between the two spectra when they are offset by *i* pixels. This takes a value between -1 and 1, where a value of 1 indicates that the spectra have exactly the same shape; a value of 0 indicates that the spectra are uncorrelated; and a value of -1 indicates that the spectra are perfectly anti-correlated (which is rather unlikely to be the case).

To estimate the RV of a spectrum, we locate the peak of its cross-correlation function with each template spectrum. The location of the peak measures the wavelength offset between the test spectrum and each template. The velocity resolution required from 4MOST translates into a requirement to determine the position of this peak at sub-pixel accuracy: in fact, with an accuracy of around 0.02 pixel.

To find the sub-pixel position of the peak of the CCF, we analytically fit a quadratic polynomial through the three CCF points which straddle the peak. We then analytically determine the position of the peak of this polynomial, as well as the value of the CCF at the peak. The value of the CCF measures the degree of fit between the template and the test spectrum, which we use to weight the RV estimates from each template.

The result from this process is a total of 225 weighted RV estimates for each test spectrum, since we cross-correlate each of the three arms of each spectrum separately with 75 template spectra. We reject any RV estimates with correlation coefficients below 0.6, which indicates a poor fit between the template and the test spectrum.

Of these 225 RV estimates, we reject the lowest and highest 25% of the RV estimates. This is important since any erroneous fits which are offset by more than a few pixels will tend to yield RV estimates close to the speed of light. These large outlying values can skew the mean of the distribution even if they have very low weights.

We then form a weighted mean of the remaining 50% of the RV estimates, weighting each RV estimate by the value of the CCF at the peak. We estimate of the uncertainty in each RV determination by computing the weighted standard deviation of the RV estimates.

5.4 Speed performance

This cross-correlation algorithm is fast and efficient. On an Intel i7 6700K desktop, cross-correlation with 75 templates utilises a single core for around 1.5 sec per test spectrum, for each of LRS and HRS. This performance scales linearly in proportion to the number of template spectra used.

6 Performance vs SNR

The histograms below show the accuracy achieved by our code in recovering the RVs of 20,000 synthetic test spectra at four different SNRs. The horizontal axis shows the absolute offset between the RV estimate generated by our code, and the actual RV that had been applied to the test spectrum.

Even at the lowest SNR tested (SNR/pixel = 20), we recover the RVs of all 20,000 HRS test spectra to within ± 1 km/s. At the same SNR, we recover the RVs of 99.6% of the LRS test spectra to within ± 2 km/s, which improves to 99.8% at SNR/pixel = 30.

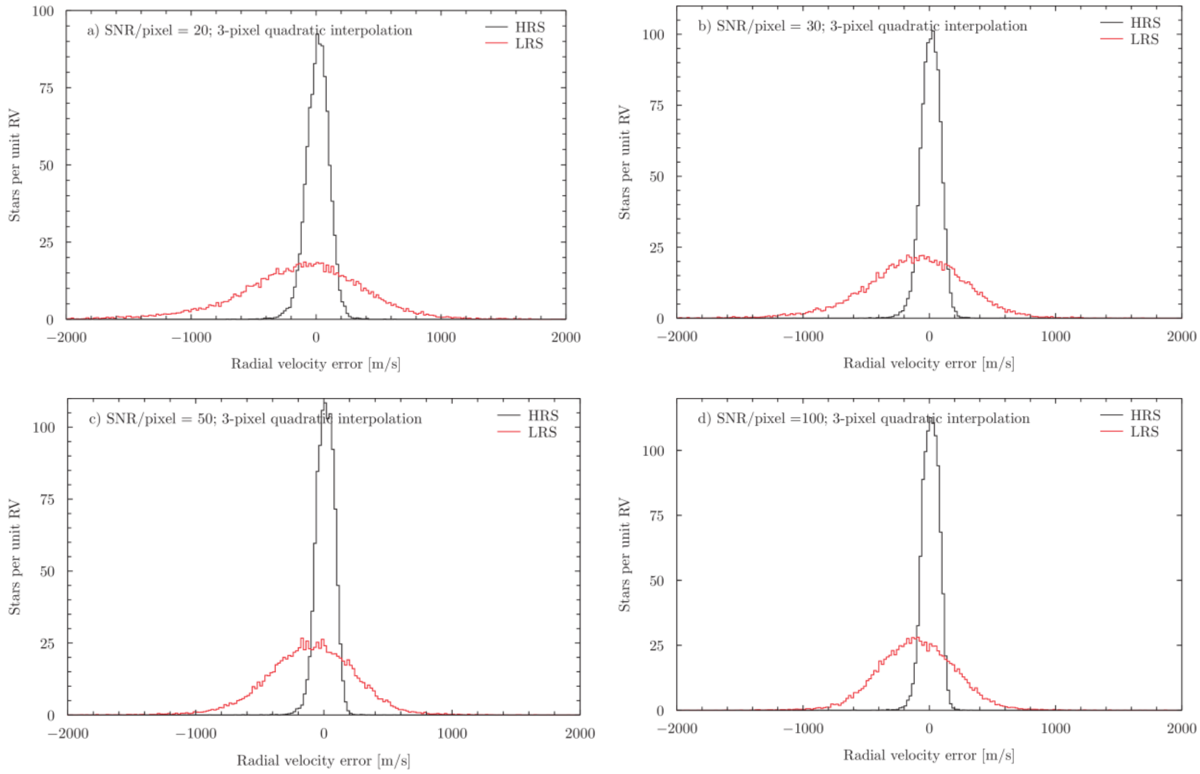


Figure 6-1: Performance vs SNR when all three HRS or LRS arms are considered together.

The table below summarises the mean and standard deviation of each offset distribution, together with the percentage of the test objects for which we were able to recover the RVs to within ±1 km/s and ±2 km/s. The mean is a measure of the systematic offset in our RV estimates, and the standard deviation is a measure of the scatter.

Mode	SNR/pixel	Mean offset [m/s]	Standard deviation [m/s]	% recovered within ± 1km/s	% recovered within ± 2km/s
HRS	20	12.2	91.2	100.0%	100.0%
	30	10.0	77.4	100.0%	100.0%
	50	10.8	68.7	100.0%	100.0%
	100	10.5	65.0	100.0%	100.0%
LRS	20	-127.3	509.6	93.7%	99.6%
	30	-120.1	405.8	97.1%	99.8%
	50	-103.3	346.4	98.6%	99.9%
	100	-98.0	305.6	99.3%	100.0%

7 Performance versus SNR

In the histograms below, we repeat this analysis, but derive RVs for each of the 20,000 test objects from each 4MOST arm individually. The 4MOST arms are numbered in ascending wavelength order, so that HRS 0 is the blue arm, and HRS 2 is the red arm.

All of our RV determinations are weighted averages of estimates obtained from each of the arms separately, and so the performance of our RV code depends upon being able to derive accurate RV estimates from each arm individually. It is therefore valuable to understand and optimise the error distribution in the estimates generated by each arm.

The histograms below show substantial systematic offsets and scatter in the estimates from some of the arms, which we seek to understand in the following sections.

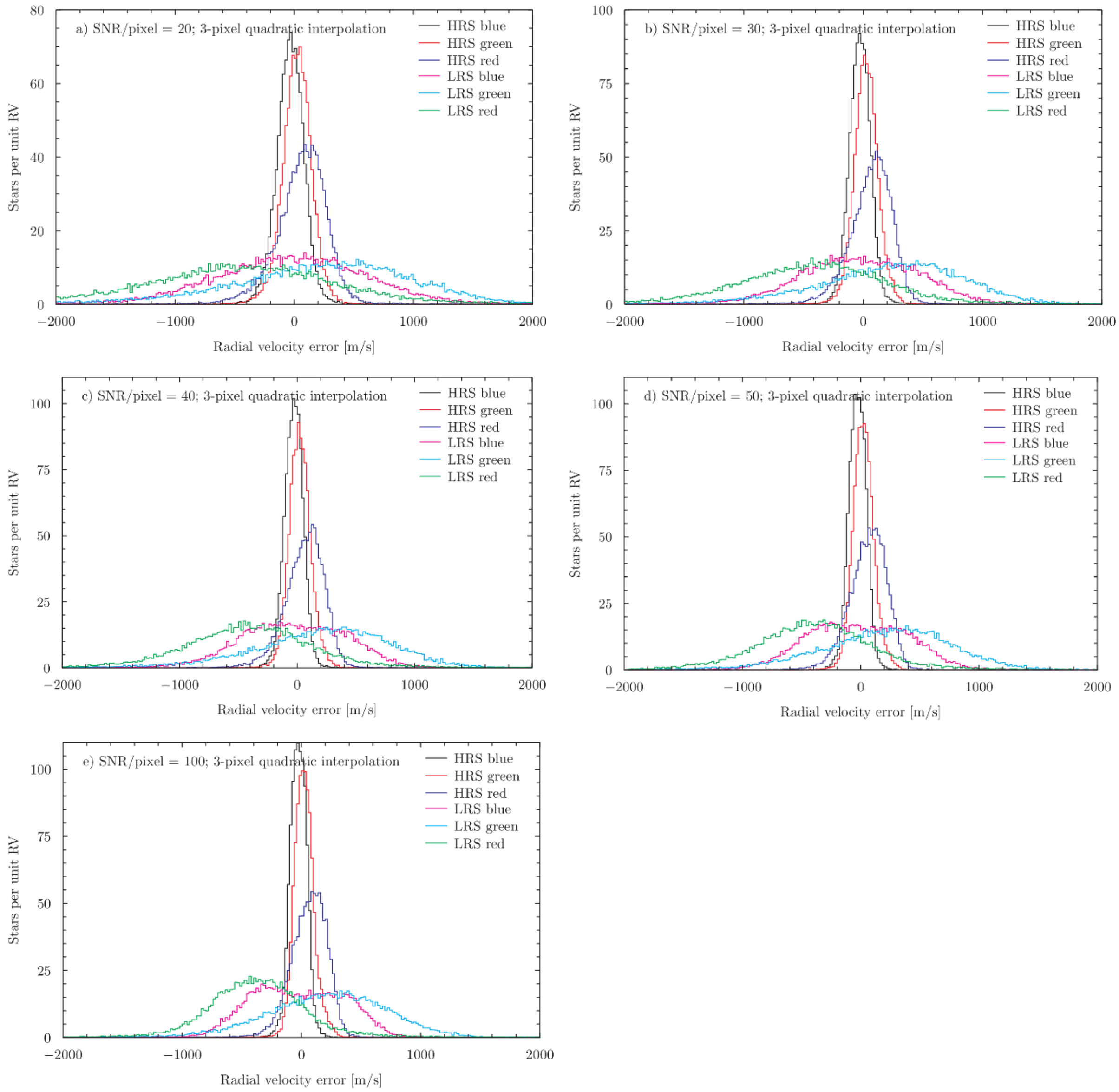


Figure 7-1: Performance vs SNR/pixel for each 4MOST arm individually.

8 Performance across the Kiel diagram

In the Kiel diagrams below, we map these RV offsets as a function of T_{eff} and $\log(g)$. The test objects are colour-coded according to the offset in their RV determinations at SNR/pixel = 50, which was configuration d in the previous sections.

Several arms show systematic RV biases as a function of T_{eff}, suggesting that our code could be improved by selectively cross-correlating each test spectrum with only templates which closely match the T_{eff} of the test object. We do not pursue this possibility any further in this report, since the RV code already comfortably meets the required accuracy for 4MOST.

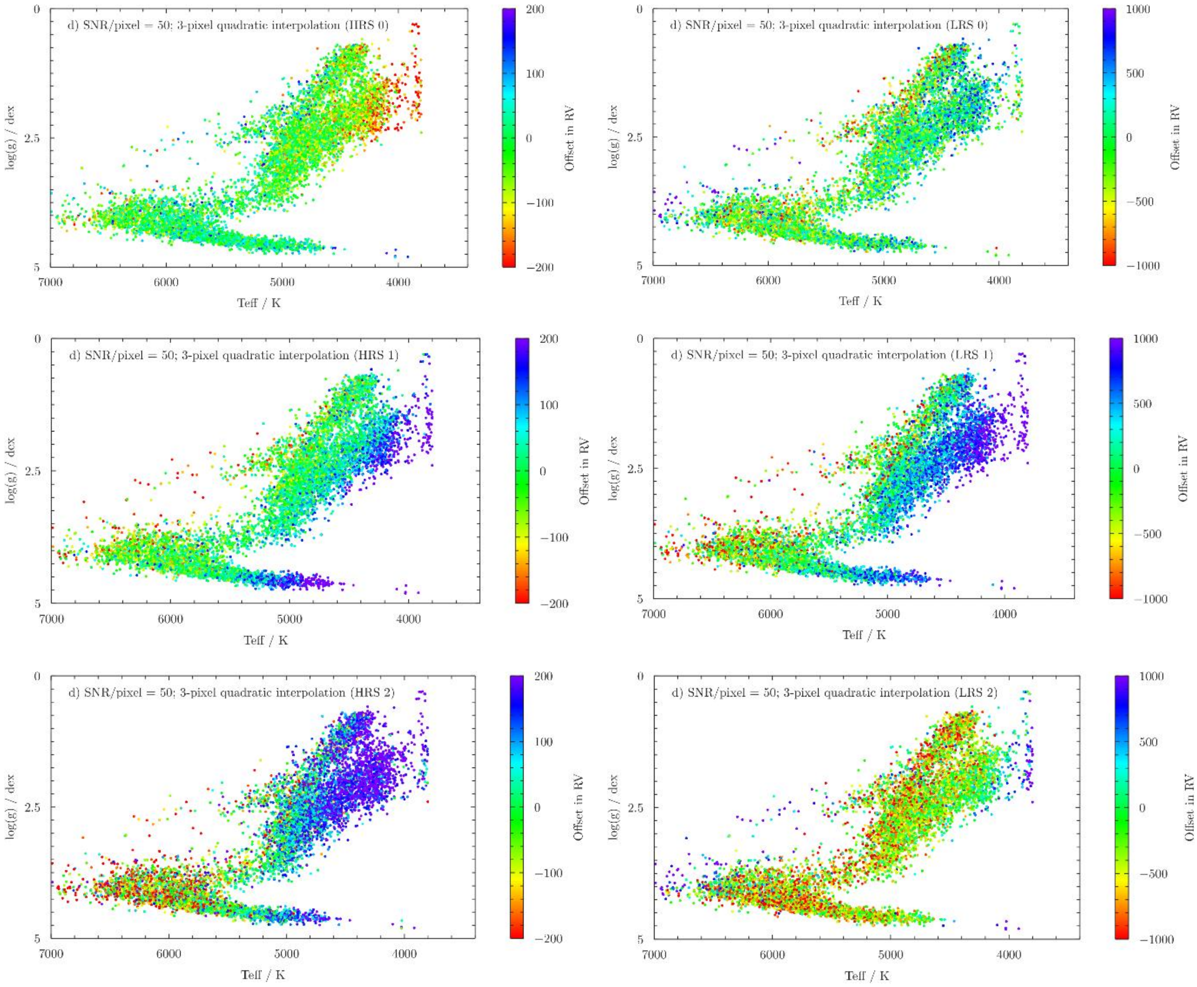


Figure 8-1: Performance across the Kiel diagram for each 4MOST arm individually.

9 RV error vs RV

Below, we plot the offsets in the RV estimates from each 4MOST arm as a function of the true value of the RV which we applied to each test object. Once again, this test was conducted at SNR/pixel = 50, which was configuration d in the previous sections.

We see that a large contribution to the errors in our RV estimates takes the form of a periodic oscillation with RV. The period of this oscillation closely matches the width of each 4MOST pixel in velocity space, which implies that it is due to an inaccuracy in the interpolation scheme we use to find the sub-pixel position of the peak of the CCF.

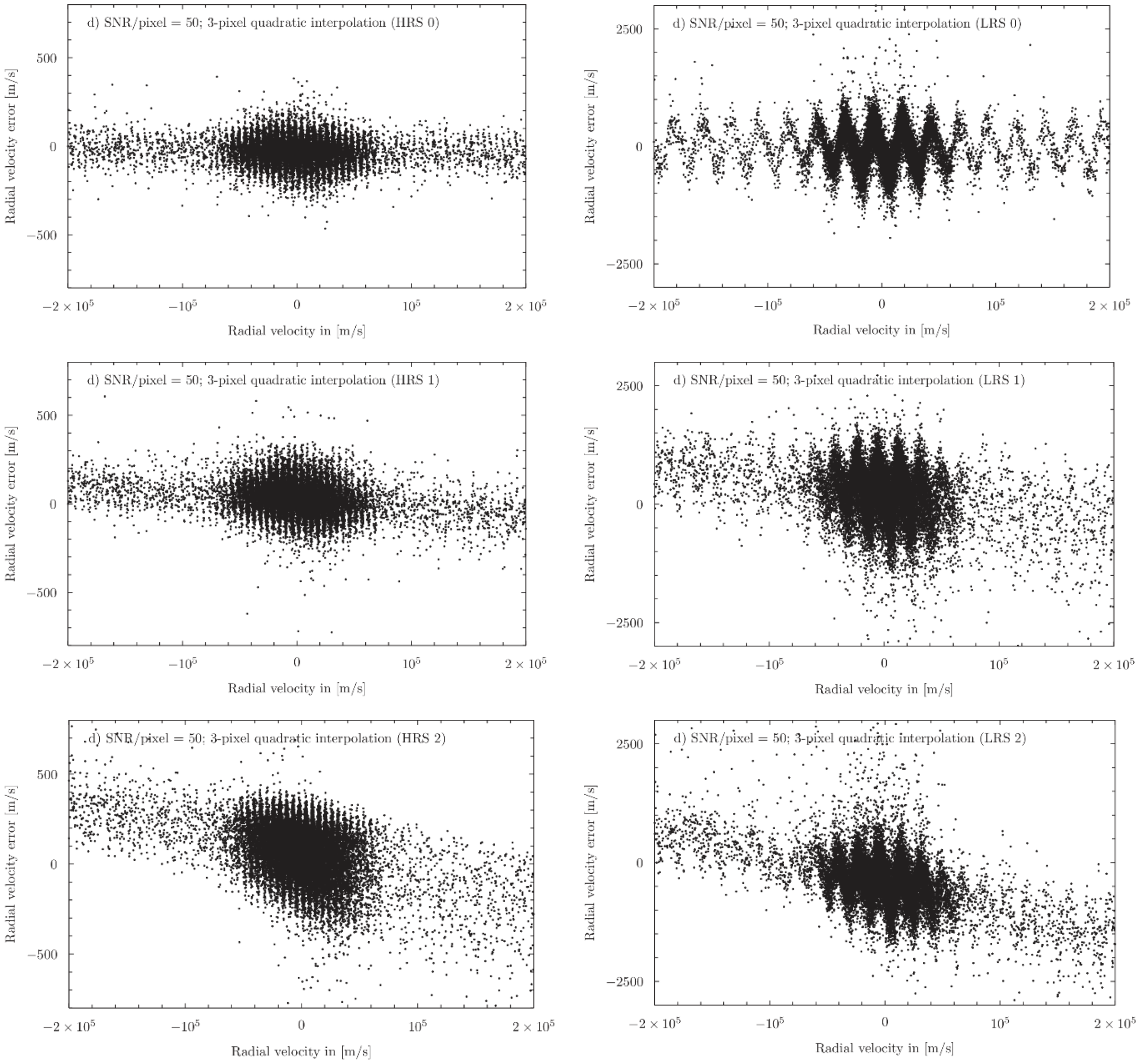


Figure 9-1: RV error as a function of RV for each 4MOST arm.

10 RV error as a function of the inter-pixel position of the CCF peak

In the plots below, we fold the RV offsets on the width of the pixels of each arm in velocity space. This verifies that the periodic oscillations above repeat with a period that equals the width of the pixels in velocity space.

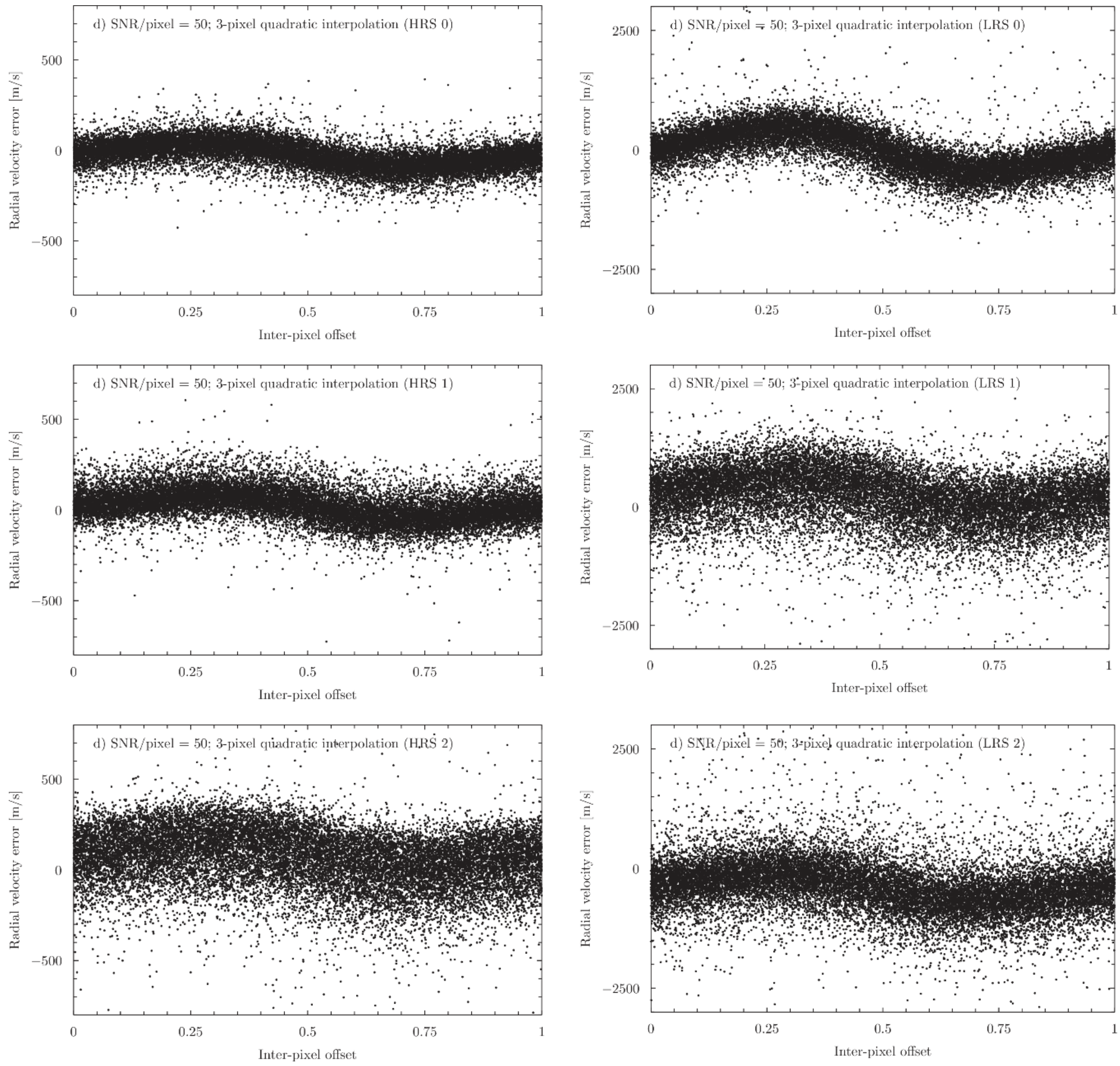


Figure 10-1: Performance as a function of the inter-pixel position of the CCF peak.

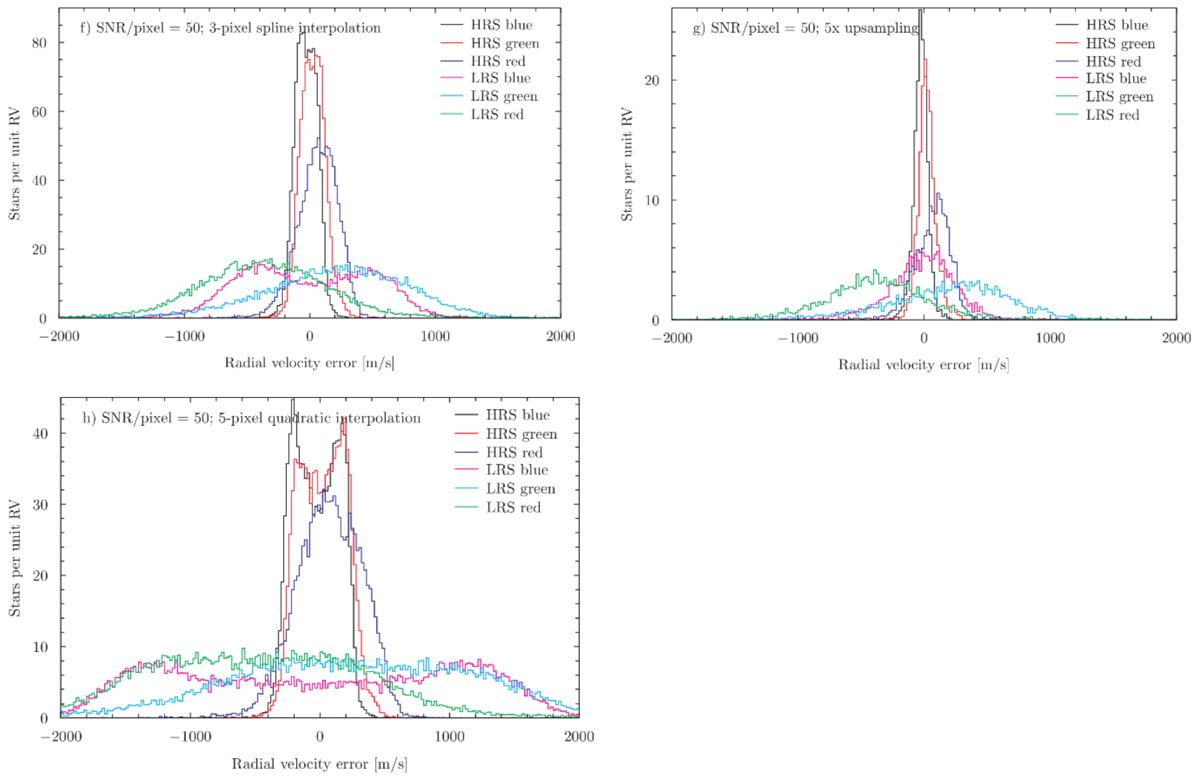
11 Alternative interpolation strategies

In this section, we investigate three alternative methods for interpolating the cross-correlation function to find the sub-pixel position of its peak. In panel (f) we fit a cubic spline through the three pixels around the peak of the CCF, instead of a quadratic polynomial. In panel (g) we up-sample the spectra by a factor of five using cubic spline interpolation prior to cross-correlation. In panel (h), we fit a quadratic polynomial through the five pixels around the peak of the CCF, rather than three pixels.

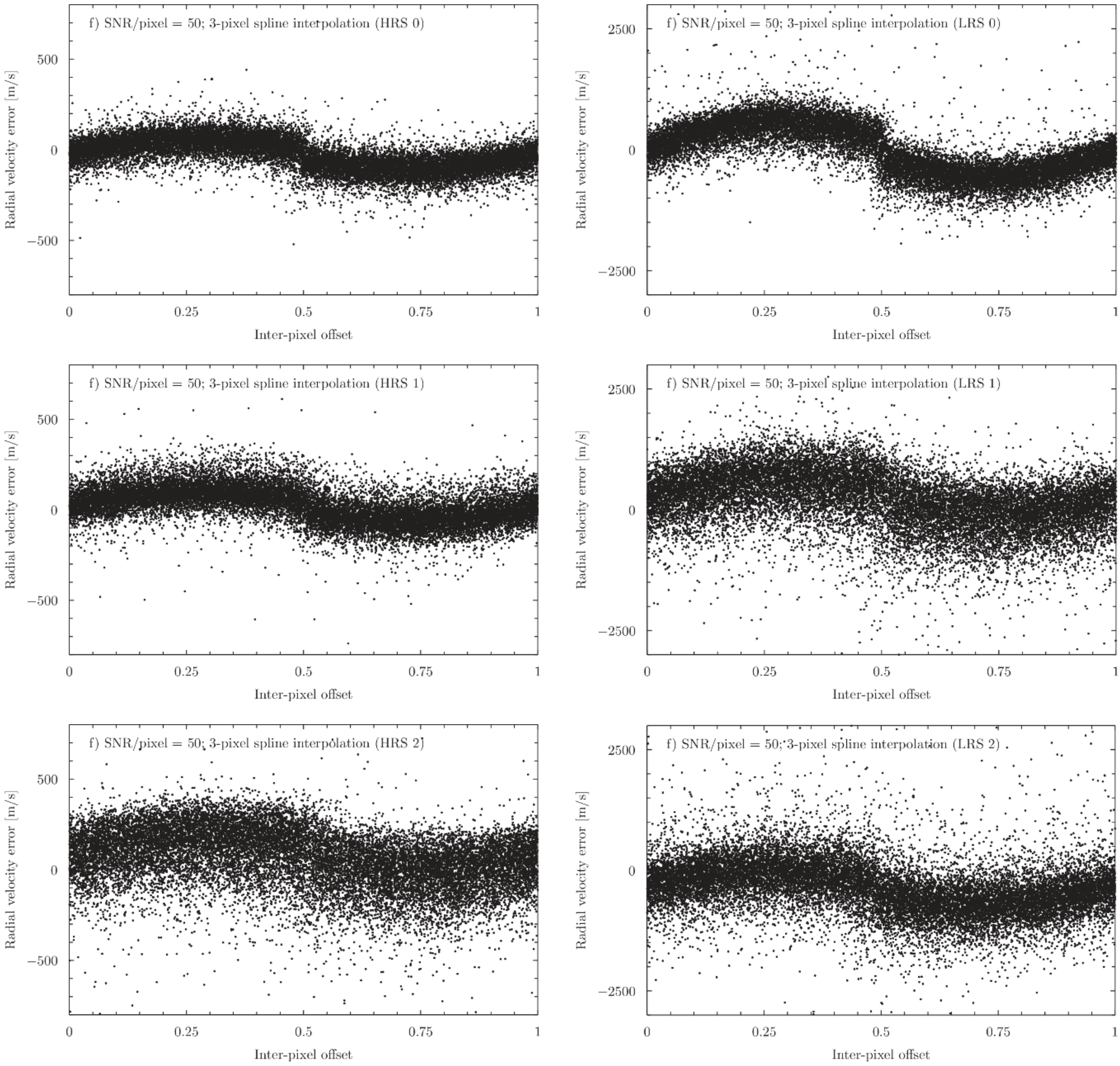
Computationally, schemes (f) and (h) are no more expensive than the fits shown previously. However, scheme (g) is significantly more expensive, since the size of spectra being cross-correlation is increased by a factor of five. Theoretically, the computational cost of cross correlation is expected to scale as $n\log(n)$. Using numpy's `correlate` cross-correlation function, we found the time required for the cross correlation increased by an order of magnitude.

To illustrate the effect that these alternative schemes have on the errors that are periodic in RV, in the following sections we plot the RV errors as a function of the inter-pixel position of the CCF peak, as previously.

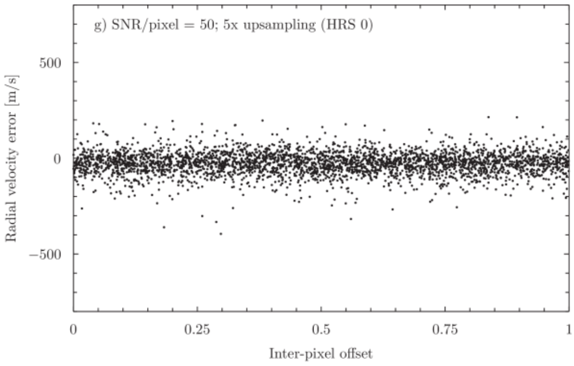
We conclude that scheme (h) performs much worse than fitting a polynomial through only three pixels. Up-sampling the spectra prior to cross-correlation (scheme g) substantially improves our results, however.

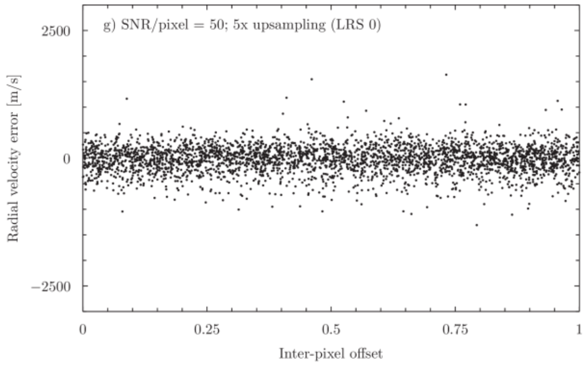


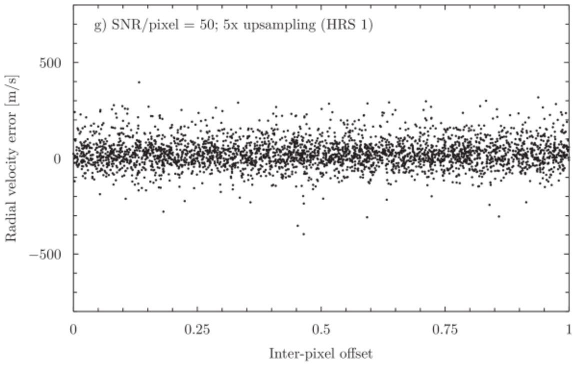
11.1 Scheme f: Three-pixel spline interpolation

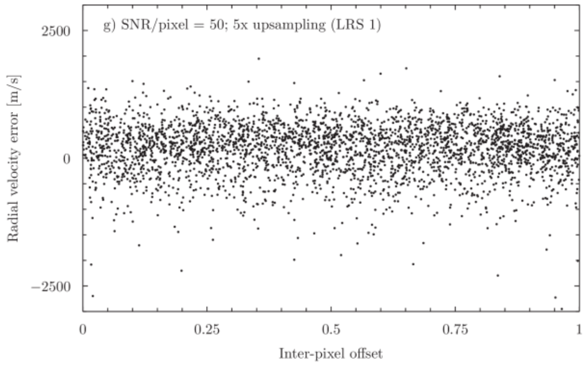


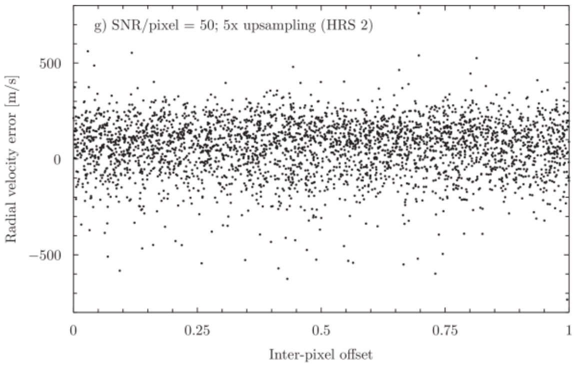
11.2 Scheme g: 5x up-sampling

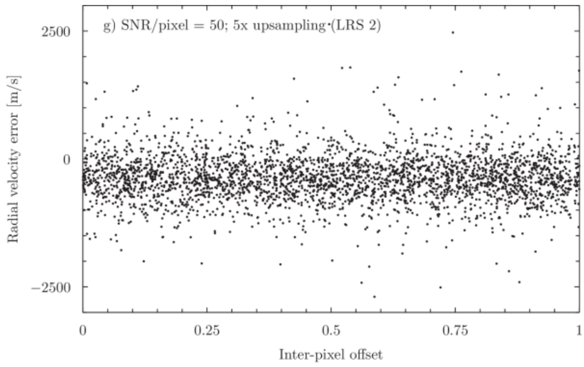




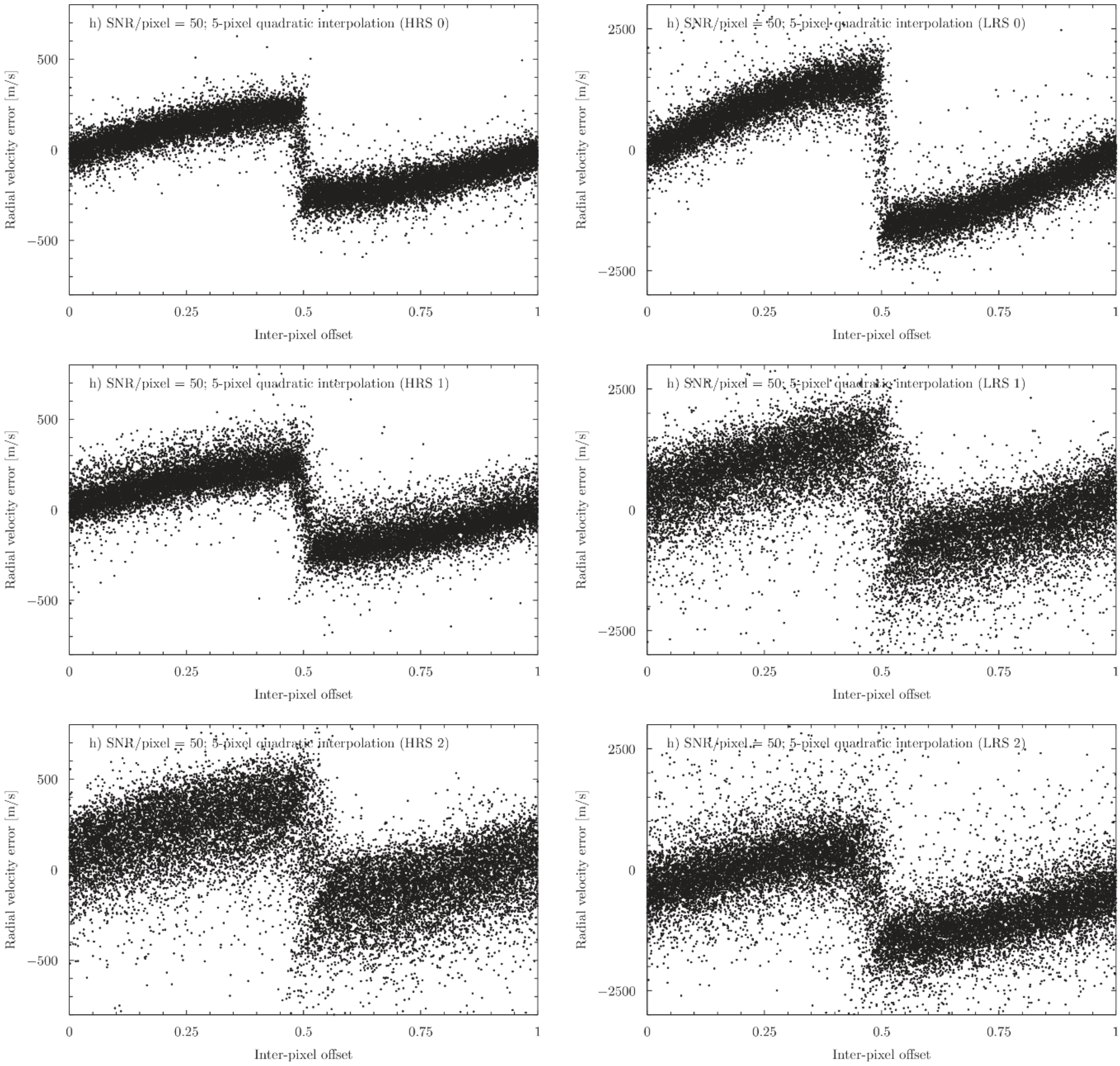








11.3 Scheme h: Five-pixel quadratic interpolation



12 Testing on real spectra

In the previous sections, we have tested our RV code exclusively on synthetic spectra. To verify that these results also extend to real observed spectra, we used the PEPSI observed spectra of the Gaia benchmark stars[1] as an alternative library of test spectra. As with the synthetic test spectra, we applied randomly chosen RVs to them, and used 4FS to reduce the resolution of spectra to match observations that we might expect 4MOST to make of these stars.

[1] See https://arxiv.org/abs/1712.06967

In our tests, we were able to recover the RVs of 99.2% of HRS observations of these stars to within 1 km/s. We recovered 99.5% of the RVs to within 2 km/s.

We were able to recover the RVs of 92.2% of LRS observations of the same stars to within 2 km/s.

13 Extending to other populations of stars

In due course, it will be necessary for 4GP to be able to determine the radial velocities of classes of stars other than the FGK stars tested here. For example, 4MOST will also observe OB stars and white dwarfs.

Our preliminary tests on the spectra of OB stars suggest that it will not be sufficient to simply extend the population of template stars used in the cross-correlation to include other types of stars. The spectra of OB stars are sufficiently different from FGK stars that their RVs cannot be estimated by cross-correlation with FGK stellar templates. Thus, it will be necessary to exclude these templates when fitting the RVs of hotter stars.

One possibility would be to place a threshold on the minimum Pearson correlation coefficient between each template and each test spectrum for that template to be used in the RV estimation process. In our initial studies, we have struggled to get this to work reliably; in our tests, around 1-2% of stars were correlated with the wrong templates and yielded RV estimates that were wrong by more than 5 km/s.

An alternative would be to create entirely separate libraries of template spectra which are appropriate for different populations of stars. Each star would be cross-correlated with all the libraries in turn. The various templates in each library will produce a wide scatter in their RV estimates if they are a poor match to the input spectrum, and so it should be easy to identify which library produced the most reliable RV estimate.

14 Conclusions

We have presented a study of the use of cross correlation to derive the radial velocities of stars from 4MOST observations.

When mock observations are made at SNR/pixel=50 of synthetic spectra which have had radial velocities applied to them, we have shown that we can recover the radial velocity to within the required precision stated in the IWG7 Management Plan [AD1] for 100% of HRS spectra, and

99.9% of LRS spectra.

We have performed an initial check that these results also apply to real observed spectra, by injecting synthetic RVs into the PEPSI observations of the Gaia Benchmark Stars. The same cross-correlation framework is again able to recover radial velocities by cross correlation to within the required precision for 99.2% of HRS spectra, and 90.2% of LRS spectra.

Appendix A List of Acronyms

List of Acronyms	
4MOST	4-metre Multi-Object Spectroscopic Telescope