The Astrophysical Journal, 812:18 (15pp), 2015 October 10 doi:10.1088/0004-637X/812/1/18

PERIODOGRAMS FOR MULTIBAND ASTRONOMICAL TIME SERIES

Jacob T. VanderPlas[1] and Željko Ivezić[2]
[1] eScience Institute, University of Washington, Seattle, WA, USA
[2] Department of Astronomy, University of Washington, Seattle, WA, USA
Received 2015 February 5; accepted 2015 August 24; published 2015 October 6

ABSTRACT

This paper introduces the *multiband periodogram*, a general extension of the well-known Lomb–Scargle approach for detecting periodic signals in time-domain data. In addition to advantages of the Lomb–Scargle method such as treatment of non-uniform sampling and heteroscedastic errors, the multiband periodogram significantly improves period finding for randomly sampled multiband light curves (e.g., Pan-STARRS, DES, and LSST). The light curves in each band are modeled as arbitrary truncated Fourier series, with the period and phase shared across all bands. The key aspect is the use of Tikhonov regularization which drives most of the variability into the so-called base model common to all bands, while fits for individual bands describe residuals relative to the base model and typically require lower-order Fourier series. This decrease in the effective model complexity is the main reason for improved performance. After a pedagogical development of the formalism of least-squares spectral analysis, which motivates the essential features of the multiband model, we use simulated light curves and randomly subsampled SDSS Stripe 82 data to demonstrate the superiority of this method compared to other methods from the literature and find that this method will be able to efficiently determine the correct period in the majority of LSST's bright RR Lyrae stars with as little as six months of LSST data, a vast improvement over the years of data reported to be required by previous studies. A Python implementation of this method, along with code to fully reproduce the results reported here, is available on GitHub.

Key words: methods: data analysis – methods: statistical – surveys

1. INTRODUCTION

Many types of variable stars show periodic flux variability (Eyer & Mowlavi 2008). Periodic variable stars are important both for testing models of stellar evolution and for using such stars as distance indicators (e.g., Cepheids and RR Lyrae stars). One of the first and main goals of the analysis is to detect variability and to estimate the period and its uncertainty. A number of parametric and non-parametric methods have been proposed to estimate the period of an astronomical time series (e.g., Graham et al. 2013 and references therein).

The most popular non-parametric method is the phase dispersion minimization (PDM) introduced by Stellingwerf (1978). Dispersion per bin is computed for binned phased light curves evaluated for a grid of trial periods. The best period minimizes the dispersion per bin. A similar and related non-parametric method that has been recently gaining popularity is the SuperSmoother routine (Reimann 1994). It uses a running mean or running linear regression on the data to fit the observations as a function of phase to a range of periods. The best period minimizes a figure-of-merit, adopted as weighted sum of absolute residuals around the running mean. Neither the SuperSmoother algorithm nor the PDM method require a priori knowledge of the light curve shape.

The most popular parametric method is the Lomb–Scargle periodogram, which is discussed in detail in Section 2. The Lomb–Scargle periodogram is related to the χ^2 for a least-square fit of a single sinusoid to data and can treat non-uniformly sampled time series with heteroscedastic measurement uncertainties. The underlying model of the Lomb–Scargle periodogram is nonlinear in frequency and so the likelihood surface in frequency is non-convex. This non-convexity is readily apparent in the many local maxima of the typical periodogram, which makes it difficult to find the maximum via standard numerical optimization routines. Thus in practice the global maximum of the periodogram is often found by a brute-force grid search (for details see, e.g., Ivezić et al. 2014).

A more general parametric method based on the use of a continuous-time autoregressive moving average (CARMA) model was recently introduced by Kelly et al. (2014). CARMA models can also treat non-uniformly sampled time series with heteroscedastic measurement uncertainties and can handle complex variability patterns.

A weakness of all these standard methods is that they require homogeneous measurements—for astronomy data, this means that successive measurements must be taken through a single photometric bandpass (filter). This has not been a major problem for past surveys because measurements are generally taken through a single photometric filter (e.g., LINEAR, Sesar et al. 2011) or nearly simultaneously in all bands at each observation (e.g., SDSS, Sesar et al. 2010). For the case of simultaneously taken multiband measurements, Süveges et al. (2012) utilized the principal component method to optimally extract the best period. Their method is essentially a multiband generalization of the well-known two-band Welch–Stetson variability index (Welch & Stetson 1993). Unfortunately, when data in each band are taken at different times, such an approach is not applicable. In such cases, past studies have generally relied on ad hoc methods such as a majority vote among multiple single-band estimates of the periodogram (e.g., Oluseyi et al. 2012).

For surveys that obtain multiband data one band at a time, such as Pan-STARRS (Kaiser et al. 2010) and DES (Flaugher 2008), and for future multicolor surveys such as LSST (Ivezić et al. 2008), this ad hoc approach is not optimal. In order to take advantage of the full information content in the available data, it would be desirable to have a

single estimate of the periodogram that accounts for all observed data in a manner independent of assumptions about the underlying spectrum of the object. We propose such a method in this paper.

The proposed method is essentially a generalization of the Lomb–Scargle method to a multiband case. The light curves in each band are modeled as an arbitrary truncated Fourier series, with the period and optionally the phase, shared across all bands. The key aspect enabling this approach is the use of a Tikhonov regularization (discussed in detail in Section 4.3) which drives most of the variability into the so-called *base model* common to all bands, while fits for individual bands describe residuals relative to the base model and typically require a lower-order Fourier series. This regularization-driven decrease in effective model complexity is the main reason for improved performance.

The remainder of the paper is organized as follows. Sections 2–4 offer a review of essential concepts in least squares modeling and least squares spectral analysis, as well as their relationship to common periodogram estimates: in Section 2 we provide a brief review of least-squares periodic fitting, in Section 3 we derive the matrix-based formalism for the single-band least squares spectral analysis used throughout the rest of this work, and Section 4 introduces several extensions and generalizations to the single-band model that the matrix formalism makes possible, including floating mean models, truncated Fourier models, and regularized models. Sections 5–7 present our new developments: in Section 5, we use the ideas and formalism of Sections 2–4 as motivation for the *multiband periodogram* and show some examples of its use on simulated data. In Section 6 we apply this method to measurements of 483 RR Lyrae stars first explored by Sesar et al. (2010, hereafter S10), and in Section 7 we explore the performance of the method for simulated observations from the LSST survey. We conclude in Section 8.

2. BRIEF OVERVIEW OF PERIODIC ANALYSIS

The detection and quantification of periodicity in time-varying signals is an important area of data analysis within modern time-domain astronomical surveys. For evenly spaced data, the *periodogram*, a term coined by Schuster (1898), gives a quantitative measure of the periodicity of data as a function of the angular frequency ω. For data $\{y_k\}_{k=1}^N$ measured at equal intervals $t_k = t_0 + k\Delta t$, Schuster's periodogram, which measures the spectral power as a function of the angular frequency, is given by

$$C(\omega) = \frac{1}{N}\left|\sum_{k=1}^{N} y_k e^{i\omega t_k}\right|^2, \tag{1}$$

and can be computed very efficiently using a fast Fourier transform.

Because astronomical observing cadences are rarely so uniform, many have looked at extending the ideas behind the periodogram to work with unevenly sampled data. Most famously, Lomb (1976) and Scargle (1982) extended earlier work to define the *normalized periodogram*:

$$P_N(\omega) = \frac{1}{2\,V_y}\left[\frac{\left[\sum_k (y_k - \bar{y})\cos\omega(t_k - \tau)\right]^2}{\sum_k \cos^2\omega(t_k - \tau)} + \frac{\left[\sum_k (y_k - \bar{y})\sin\omega(t_k - \tau)\right]^2}{\sum_k \sin^2\omega(t_k - \tau)}\right], \tag{2}$$

where $\bar{y}$ is the mean, V_y is the variance of the data $\{y_k\}$, and τ is the time-offset which orthogonalizes the model and makes $P_N(\omega)$ independent of a translation in t (see Press et al. 2007 for an in-depth discussion). Lomb (1976) showed that this time-offset has a deeper effect: namely, it gives P_N a form similar to previous extensions of $C(\omega)$, while leaving P_N identical to the estimate of harmonic content given a least-squares fit to a single-component sinusoidal model,

$$d(t) = A\,\sin(\omega t + \phi). \tag{3}$$

This long-recognized connection between spectral power and least-squares fitting methods was solidified by Jaynes (1987, pp. 1–37), who demonstrated that the least-squares periodogram method is a sufficient statistic for inferences about a stationary frequency signal in the presence of Gaussian noise. Building on this result, Bretthorst (1988) explored the extension of these methods to more complicated models with multiple frequency terms, non-stationary frequencies, and other more sophisticated models within a Bayesian framework.

While the important features of least squares frequency estimation via Lomb–Scargle periodograms have been discussed elsewhere, we will present a brief introduction to the subject in the following section. In particular, we re-express the problem in a matrix-based formalism that makes clear how the basic approach motivated by Lomb (1976), Scargle (1982), and others can be extended to more sophisticated models including the multiband periodogram proposed in this work.

3. STANDARD LEAST-SQUARES SPECTRAL FITTING

In this section we present a brief quantitative introduction to the least-squares fitting formulation of the normalized periodogram of Equation (2). We denote N observed data points as

$$D = \{t_k, y_k, \sigma_k\}_{k=1}^N \tag{4}$$

where t_k is the time of observation, y_k is the observed value (typically a magnitude), and σ_k describes the Gaussian errors on each value. For notational simplicity we will assume without loss of generality that the data y_k are centered such that the measurements within each band satisfy

$$\frac{\sum_k w_k y_k}{\sum_k w_k} = 0 \tag{5}$$

where the weights are $w_k = \sigma_k^{-2}$. Though this assumption is essential to the simpler models presented in this section, it will become superfluous with the floating-mean models described in later sections.

3.1. Stationary Sinusoid Model

The normalized periodogram of Equation (2) can be derived from the normalized χ^2 of the best-fit single-term stationary sinusoidal model given in Equation (3). To make the problem linear, we can re-express the model in terms of the parameter vector $\theta = [A \cos \phi, A \sin \phi]$ so that our model is

$$y(t|\omega, \theta) = \theta_1 \sin(\omega t) + \theta_2 \cos(\omega t). \tag{6}$$

For a given ω, the maximum likelihood estimate of the parameters θ can be found by minimizing the χ^2 of the model, which is given by

$$\chi^2(\omega) = \sum_k \frac{\left[y_k - y\left(t_k|\omega, \theta\right)\right]^2}{\sigma_k^2}. \tag{7}$$

For the single-term Fourier model, it can be shown (see, e.g., Ivezić et al. 2014) that

$$\chi^2_{\min}(\omega) = \chi_0^2\left[1 - P_N(\omega)\right] \tag{8}$$

where $P_N(\omega)$ is the normalized periodogram given in Equation (2)[3] and χ_0^2 is the reference χ^2 for a constant model, which due to the assumption in Equation (5) is simply $\chi_0^2 = \sum_k (y_k/\sigma_k)^2$.

3.2. Matrix Formalism

A standard way of compactly expressing least-squares models is via matrix expressions (see, e.g., Brandt 1970). Likewise, the expressions related to the stationary sinusoid model can be expressed more compactly by defining the following matrices:

$$X_\omega = \begin{bmatrix} \sin(\omega t_1) & \cos(\omega t_1) \\ \sin(\omega t_2) & \cos(\omega t_2) \\ \vdots & \vdots \\ \sin(\omega t_N) & \cos(\omega t_N) \end{bmatrix};$$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}; \quad \Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_N^2 \end{bmatrix}. \tag{9}$$

With these definitions, the model in Equation (6) can be expressed as a simple linear product, $y(t|\omega, \theta) = X_\omega \theta$, and the model and reference χ^2 can be written as

$$\chi^2(\omega) = \left(y - X_\omega \theta\right)^T \Sigma^{-1} \left(y - X_\omega \theta\right) \tag{10}$$

$$\chi_0^2 = y^T \Sigma^{-1} y. \tag{11}$$

The expression for the normalized periodogram can be computed by finding via standard methods the value of θ that minimizes $\chi^2(\omega)$ and plugging the result into Equation (8). This yields

$$P_N(\omega) = \frac{y^T \Sigma^{-1} X_\omega \left[X_\omega^T \Sigma^{-1} X_\omega\right]^{-1} X_\omega^T \Sigma^{-1} y}{y^T \Sigma^{-1} y}. \tag{12}$$

We note that Equation (12) is equivalent to Equation (2) in the homoscedastic case with $\Sigma \propto V_y I$.[4]

3.3. Simple Single-band Period Finding

As an example of the standard periodogram in action, we perform a simple single-band harmonic analysis of simulated r-band observations of an RR Lyrae light curve based on empirical templates derived in S10 (Figure 1). The observations are of a star with a period of 0.622 days and take place on 60 random nights over a six-month period, as seen in the left panel.

The upper-right panel shows the normalized periodogram for this source as a function of period. While the power does peak at the true period of 0.622 days, an aliasing effect is readily apparent near $P = 0.38$. This additional peak is due to the beat frequency between the true period P and the observing cadence of $\sim$1 day. This beat frequency is the first in a large sequence: for nightly observations, we would expect to find excess power at periods $P_n = P/(1 + nP)$ days for any integer n. The strong alias in Figure 1 corresponds to the $n = 1$ beat period $P_n = 0.383$. Though it is possible to carefully correct for such aliasing by iteratively removing contributions from the estimated window function (e.g., Roberts et al. 1987), we will ignore this detail in the current work.

The lower-right panel of Figure 1 shows the maximum likelihood interpretation of this periodogram: it is a measure of the normalized χ^2 for a single-term sinusoidal model. Here we visualize the data from the left panel, but folded as a function of phase and overplotted with the best-fit single-term model. This visualization makes it apparent that the single-term model is highly biased: RR Lyrae light curves are, in general, much more complicated than a simple sinusoid. Nevertheless, the simplistic sinusoidal model is able to recover the correct frequency to a high degree of accuracy (roughly related to the width of the peak) and significance (roughly related to the height of the peak; see Scargle 1982 for details). For a more complete introduction to and discussion of the single-term normalized periodogram, refer to, e.g., Bretthorst (1988) or Ivezić et al. (2014).

4. GENERALIZING THE PERIODOGRAM MODEL

We have shown two forms of the classic normalized periodogram: Equations (2) and (12). Though the two expressions are equivalent, they differ in their utility. Because the expression in Equation (2) avoids the explicit construction of a matrix, it can be computed very efficiently. Furthermore, through clever use of the fast Fourier transform, expressions of the form of Equation (2) can be evaluated exactly for N frequencies in $\mathcal{O}[\log N]$ time (Press & Rybicki 1989).

[3] An important feature of the Lomb–Scargle approach is the modification of the model with time-offset τ tuned to orthogonalize the harmonic basis across the irregular times $\{t_i\}$. This orthogonalization cancels cross-terms in the expression of χ^2, greatly reducing the complexity of computing P_N. As discussed in footnote 4, however, this orthogonalization does not change the resulting periodogram, so it can safely be ignored for the purposes of this work.

[4] For direct comparison to the Lomb–Scargle approach, we need the equivalent of the τ parameter which orthogonalizes the basis across the observed times $\{t_i\}$. Such an orthogonalization is accomplished via the transformations $X_\omega \to X_\omega V_\omega$ and $\theta \to V_\omega^T \theta$, where V_ω is the orthogonal matrix of eigenvectors of the covariance $X_\omega^T \Sigma^{-1} X_\omega$. The V_ω terms straightforwardly cancel out of Equations (10)–(12) and the results of this section are unchanged. The general matrix formalism used here makes clear that this result applies to all the periodogram extensions mentioned in this work.

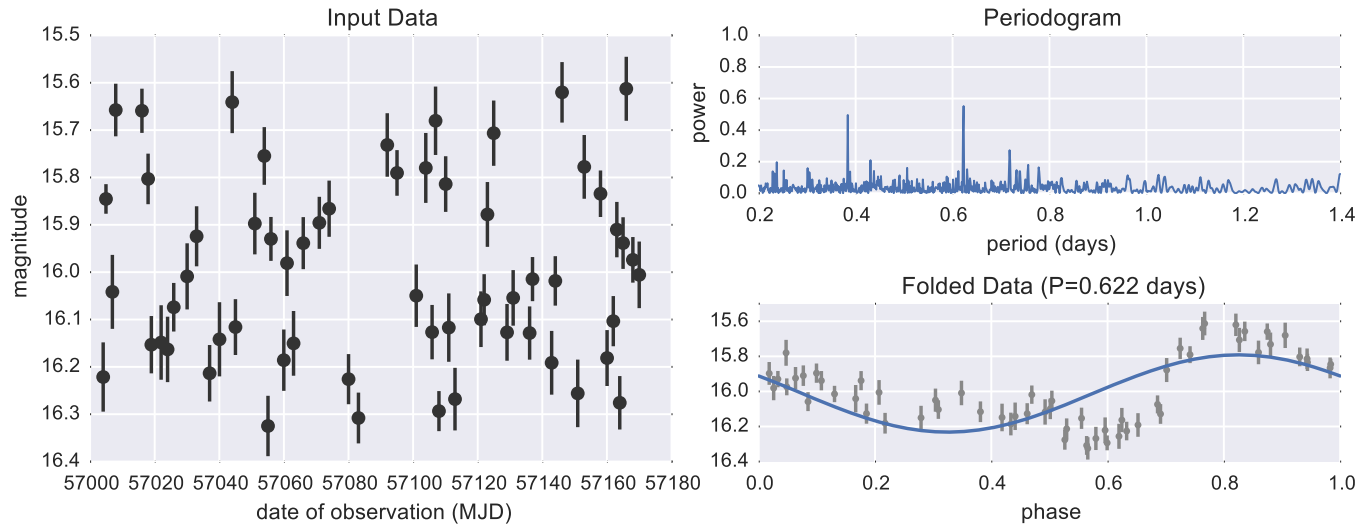

Figure 1. Illustration of the basic periodogram and its relationship to the single-term sinusoid model. The left panel shows the input data, while the right panels show the fit derived from the data. The upper-right panel shows the periodogram with a clear peak at the true period of 0.622 days, and the bottom-right panel shows the data as a function of the phase associated with this period. Note in the periodogram the presence of the typical aliasing effect, with power located at beat frequencies between the true period and the one-day observing cadence (see Section 3.3 for further discussion).

The matrix-based formulation of Equation (12), though slower than the Fourier-derived formulation, is a more general expression and allows several advantages.

1. It is straightforwardly extended to heteroscedastic and/or correlated measurement noise in the data y_k through appropriate modification of the *noise covariance matrix* Σ.
2. It is straightforwardly extended to more sophisticated linear models by appropriately modifying the *design matrix* X_ω.
3. It is straightforwardly extended to include Tikhonov/L2-regularization terms (see Section 4.3 for more details) by adding an appropriate diagonal term to the *normal matrix* $X_\omega^T \Sigma^{-1} X_\omega$.

In the remainder of this section, we will explore a few of these modifications and how they affect the periodogram and resulting model fits.

4.1. Stationary Sinusoid with Floating Mean

As an example of one of these generalizations, we will consider what has variously been called the *date-compensated discrete Fourier transform* (Ferraz-Mello 1981), the *floating-mean periodogram* (Cumming et al. 1999), and the *generalized Lomb–Scargle method* (Zechmeister & Kürster 2009). Here we use the term *floating-mean periodogram*. This method adjusts the classic normalized periodogram by fitting the mean of the model alongside the amplitudes:

$$y(t \mid \omega, \theta) = \theta_0 + \theta_1 \sin \omega t + \theta_2 \cos \omega t. \tag{13}$$

The periodogram derived from this model can be more accurate than the standard pre-centered periodogram for certain observing cadences and selection functions, and especially when searching for long-period varaibility or working with very few samples (Cumming et al. 1999). Zechmeister & Kürster (2009) detail the required modifications to the orthogonalized harmonic formalism of Equation (2) to allow the mean to float in the model. In the matrix formalism, the modification is much more straightforward: all that is required is to add a column of ones to the X_ω matrix before computing the power via Equation (12). This column of ones corresponds to a third entry in the parameter vector θ, and acts as a uniform constant offset for all data points.

For well-sampled data, there is usually very little difference between a standard periodogram on pre-centered data and a floating-mean periodogram. Where this difference becomes important is if selection effects or observing cadences cause there to be preferentially more observations at certain phases of the light curve: a toy example demonstrating this situation is shown in Figure 2. The data are drawn from a sinusoid with Gaussian errors, and data with magnitudes fainter than 16 are removed to simulate an observational bias (left panel). Because of this observational bias, the mean of the observed data is a poor predictor of the true mean, causing the standard method to poorly fit the data and miss the input period (upper-right panel). The floating-mean approach is able to automatically adjust for this bias, resulting in a periodogram that readily detects the input period of 0.622 days (lower-right panel).

4.2. Truncated Fourier Models

As mentioned above, the standard periodogram is equivalent to fitting a single-term stationary sinusoidal model to the data. A natural extension is to instead use a multiple-term sinusoidal model with frequencies at integer multiples of the fundamental frequency (see, e.g., Bretthorst 1988). With N Fourier terms, there are $2N + 1$ free parameters, and the model is given by

$$y(t|\omega, \theta) = \theta_0 + \sum_{n=1}^{N} \left[\theta_{2n-1} \sin(n\omega t) + \theta_{2n} \cos(n\omega t)\right]. \tag{14}$$

Because this model remains linear in the parameters θ, it can be easily accommodated into the matrix formalism of Section 3.2. For example, a two-term ($N = 2$) floating-mean model can be constructed by building a design matrix X_ω with $2N + 1 = 5$

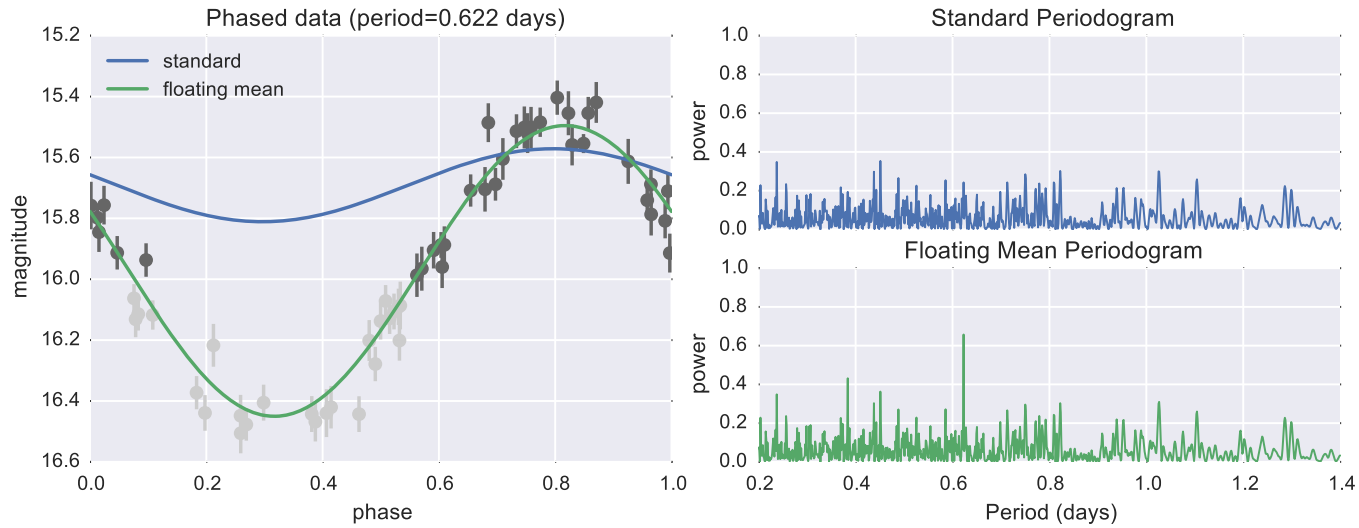

Figure 2. Illustration of the effect of the floating-mean model for censored data. The data consist of 80 observations drawn from a sinusoidal model. To mimic a potentially damaging selection effect, all observations with magnitudes fainter than 16 are removed (indicated by the light gray points). The standard and floating-mean periodograms are computed from the remaining data; these fits are shown over the data in the left panel. Because of this biased observing pattern, the mean of the observed data is a biased estimator of the true mean. The standard fixed-mean model in this case fails to recover the true period of 0.622 days, while the floating-mean model still finds the correct period.

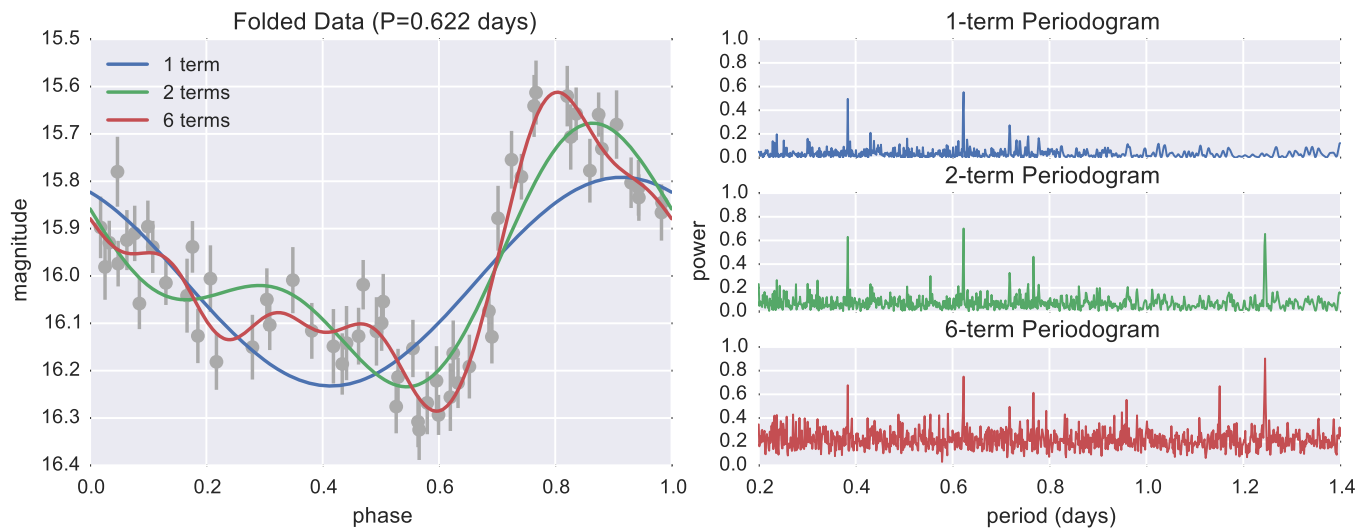

Figure 3. Model fits and periodograms for several truncated Fourier models. The data are the same as those in Figure 1. Note that in addition to the previously seen 0.38 day alias, the higher-order models will generally show periodogram peaks at multiples of the true fundamental frequency P_0: this is because for integer n less than the number of Fourier terms in the model, P_0 is a higher harmonic of the model at $P = nP_0$. Additionally, the increased degrees of freedom in the higher-order models let them fit better at any frequency, which drives up the "background" level in the periodogram.

columns:

$$X_\omega^{(2)} = \begin{bmatrix} 1 & \sin(\omega t_1) & \cos(\omega t_1) & \sin(2\omega t_1) & \cos(2\omega t_1) \\ 1 & \sin(\omega t_2) & \cos(\omega t_2) & \sin(2\omega t_2) & \cos(2\omega t_2) \\ 1 & \sin(\omega t_3) & \cos(\omega t_3) & \sin(2\omega t_3) & \cos(2\omega t_3) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \sin(\omega t_N) & \cos(\omega t_N) & \sin(2\omega t_N) & \cos(2\omega t_N) \end{bmatrix}. \tag{15}$$

Computing the power via Equation (12) using $X_\omega^{(2)}$ will give the two-term periodogram. For larger N, more columns are added, but the periodogram can be computed in the same manner. Figure 3 shows a few examples of this multiterm Fourier approach as applied to the simulated RR Lyrae light curve from Figure 1 and illustrates several important insights into the subtleties of this type of multiterm fit.

First, we see in the right panel that all three models show a clear signal at the true period of $P_0 = 0.622$ days. The higher-order models, however, also show a spike in power at $P_1 = 2P_0$: the reason for this is that for a model with more than one frequency term, the period P_0 is the first harmonic of a model with fundamental frequency $2P_0$, and the higher-order models contain the single-period result.

Second, notice that as the number of terms is increased, the general "background" level of the periodogram increases. This is due to the fact that the periodogram power is inversely related to the χ^2 of the fit at each frequency. A more flexible higher-order model can better fit the data at all periods, not just the true period. Thus in general, the observed power of a

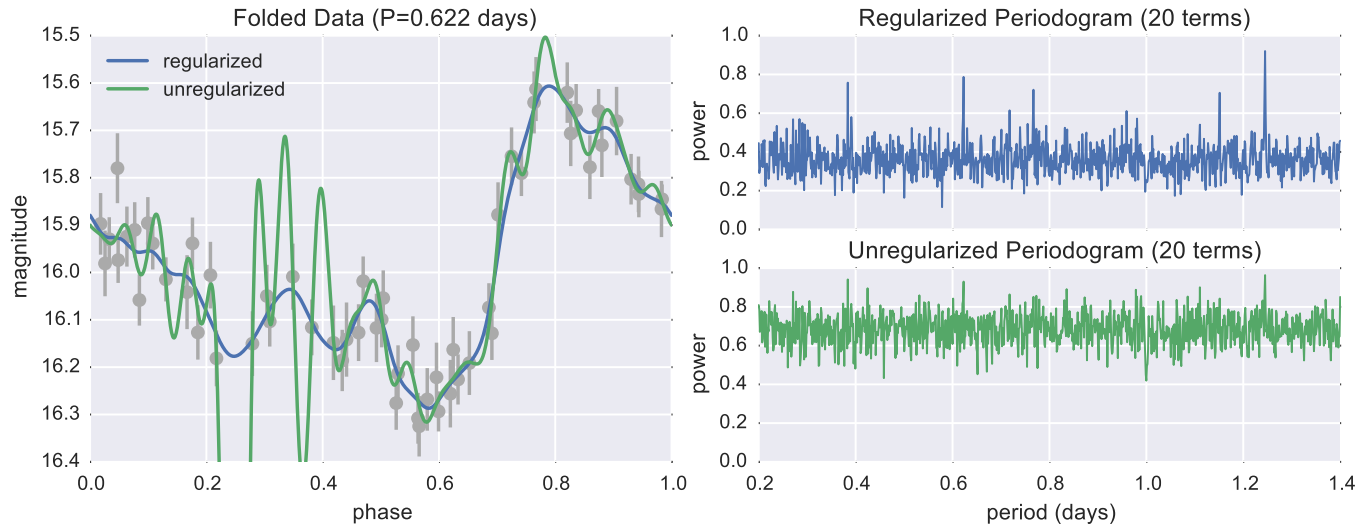

Figure 4. Effect of regularization on a high-order model. The data are the same as those in Figure 1. We fit a twenty-term truncated Fourier model to the data with and without a regularization term. Without regularization, the model oscillates widely to fit the noise in the data. The regularization term effectively damps the higher-order Fourier modes and removes this oscillating behavior, leading to a more robust model with stronger periodogram peaks.

higher-order Fourier model will be higher than the power of a lower-order Fourier model everywhere.

One might hope that when adding terms, the correct-period model would show more of an improvement than the incorrect-period model (and thus the periodogram maximum would become more pronounced in comparison to the background), but this does not generally hold. Consider that in the extreme limit in which the number model parameters is equal to the number of data points, the model has enough flexibility to fit the data perfectly at *every* frequency, and the resulting periodogram would be unity everywhere! This can only be the case if, on average, the addition of terms preferentially boosts the background level.

4.3. Regularized Models

The previous sections raise the question: How complicated a model should we use? We have seen that as we add more terms to the fit, the model will more closely describe the observed data. For very high-order models, however, such a close fit *over-fits* the data: that is, the fit is more responsive to statistical noise in the observations than to the underlying signal. This can be addressed by explicitly truncating the series at some number of terms, but we can also use a *regularization* term to mathematically enforce model simplicity.

A regularization term is an explicit penalty on the magnitude of the model parameters θ and can take a number of forms. For computational simplicity here we will use an *L2 regularization* —also known as a Tikhonov regularization (Tikhonov 1963) or Ridge regression (Hoerl & Kennard 1970)—which is a quadratic penalty term in the model parameters added to the χ^2. Mathematically, this is equivalent to using a zero-mean Gaussian prior on the model parameters in the Bayesian framework.

We encode our regularization in the matrix $\Lambda = \mathrm{diag}([\lambda_1, \lambda_2 \cdots \lambda_M])$ for a model with M parameters and construct a "regularized" χ^2:

$$\chi^2_\Lambda(\omega) = \left(y - X_\omega\theta\right)^T \Sigma^{-1}\left(y - X_\omega\theta\right) + \theta^T\Lambda\theta. \quad (16)$$

Minimizing this regularized χ^2, solving for θ, and plugging into the expression for P_N gives us the regularized counterpart of Equation (12):

$$P_{N,\Lambda}(\omega) = \frac{y^T\Sigma^{-1}X_\omega\left[X_\omega^T\Sigma^{-1}X_\omega + \Lambda\right]^{-1}X_\omega^T\Sigma^{-1}y}{y^T\Sigma^{-1}y}. \quad (17)$$

Notice that the effect of this regularization term is to add a diagonal penalty to the normal matrix $X_\omega^T\Sigma^{-1}X_\omega$, which has the additional feature that it can correct ill-posed models where the normal matrix is non-invertible. This feature of the regularization will become important for the multiband models discussed below.

In Figure 4, we compare regularized and unregularized twenty-term truncated Fourier models of our simulated RR Lyrae light curve. We use $\lambda = 0$ for the offset term and make the penalty λ_j progressively larger for each harmonic component. The regularization prevents overfitting (left panel) and results in more prominent periodogram peaks (right panel).

5. A MULTIPLE-BAND MODEL

In this section we will combine the ideas of the previous sections to construct a *multiband periodogram* which flexibly accounts for heterogeneous sources of data for a single object. To start with, we can consider one of two naïve approaches to the multi-band problem.

First, we might ignore band labels entirely and simply compute a single standard Lomb–Scargle periodogram over the full data set. This amounts to the assumption that one global model suitably fits each band and in practice will perform poorly due to the astrophysical variability between bands: in other words, the model is too simple and under-fits the data.

Second, we might treat each band entirely independently and compute a standard Lomb–Scargle periodogram on each, then use the sum of χ^2 along with Equation (8) to construct a multiband periodogram. This amounts to the assumption that the bands have completely independent phases and amplitudes and has too many free parameters to be useful in most cases of

interest. In other words, the model is too complex and over-fits the data (see Section 5.1 for further discussion).

To compute a periodogram that strikes a balance between these two extremes, we will take advantage of the easy extensibility of the matrix formalism which led to our generalizations above. The multiband model presented here contains the following features.

1. An $N_{\rm base}$-term truncated Fourier "base model" that models the shared variability among all K observed bands.
2. A set of $N_{\rm band}$-term truncated Fourier fits, each of which models the residual within a single band from the shared variability accounted for in the base model.

The total number of parameters for K bands is then $M_K = (2N_{\rm base} + 1) + K(2N_{\rm band} + 1)$. As a result, for each band k we have the following model of the observed magnitudes:

$$y_k(t|\omega, \theta) = \theta_0 + \sum_{n=1}^{N_{\rm base}} \left[\theta_{2n-1}\sin(n\omega t) + \theta_{2n}\cos(n\omega t)\right] + \theta_0^{(k)} + \sum_{n=1}^{N_{\rm band}} \left[\theta_{2n-1}^{(k)}\sin(n\omega t) + \theta_{2n}^{(k)}\cos(n\omega t)\right]. \tag{18}$$

The important feature of this model is that *all bands* share the same base parameters θ, while their offsets θ^k are determined individually. Note the potential for confusion: $N_{\rm band}$ here is not the number of observed bands, but the number of Fourier components fit to the residuals in each of the K observed bands.

We can construct the normalized periodogram for this model by building a sparse design matrix with M_K columns. Each row corresponds to a single observation through a single band. Columns corresponding to the base model and the matching observation band will have nonzero entries; all other columns will be filled with zeros. For example, the $(N_{\rm base}, N_{\rm band}) = (1, 0)$ model corresponds to one with a simple single-term periodic base frequency, and an independent constant offset term in each band. The associated design matrix depends on the particular data, but will look similar to this:

$$X_\omega^{(1,0)} = \begin{bmatrix} 1 & \sin(\omega t_1) & \cos(\omega t_1) & 1 & 0 & 0 & 0 & 0 \\ 1 & \sin(\omega t_2) & \cos(\omega t_2) & 0 & 0 & 0 & 0 & 1 \\ 1 & \sin(\omega t_3) & \cos(\omega t_3) & 0 & 0 & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & & & \vdots & & \\ 1 & \sin(\omega t_N) & \cos(\omega t_N) & 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \tag{19}$$

Here the nonzero entries of the final five columns are binary flags indicating the (u, g, r, i, z)-band of the given observation: for this example, the first row is a u-band measurement, the second is a z-band, the third is an i-band, etc., as indicated by the position of the nonzero matrix element within the row.

Upon examination of the above matrix, it is clear that the columns are not linearly independent (i.e., X_ω is low-rank), and thus the parameters of the best-fit model will be degenerate. Intuitively, this is due to the fact that if we add an overall offset to the base model, this can be perfectly accounted for by subtracting that same offset from each residual model. Mathematically, the result of this is that the normal matrix $X_\omega^T \Sigma^{-1} X_\omega$ will be non-invertible, and thus the periodogram is ill-defined. In order to proceed, then, we will either have to use a different model or use a cleverly constructed regularization term on one of the offending parameters.

We will choose the latter here and regularize all the band columns while leaving the base columns un-regularized: for the above X_ω matrix, this regularization will look like

$$\Lambda^{(1,0)} = \mathrm{diag}([0, 0, 0, \lambda, \lambda, \lambda, \lambda, \lambda]) \tag{20}$$

where λ controls the degree of regularization. As λ grows large, the model will preferentially push power into the base terms while minimizing the deviations of the model for each individual band.

Here we will choose λ to be some small fraction of the trace of the normal matrix $[X_\omega^T \Sigma^{-1} X_\omega]$. This choice ensures the multiband periodogram is well-defined, while maintaining the flexibility of the model in accounting for independent band-to-band variation. With this regularization in place, the model is well-posed and Equation (17) can be used to straightforwardly compute the power. The effective number of free parameters for such a regularized $(N_{\rm base}, N_{\rm band})$ model with K filters is $M_K^{\rm eff} = 2N_{\rm base}^{\rm eff} + K(2N_{\rm band} + 1)$ where $N_{\rm base}^{\rm eff} = \max(0, N_{\rm base} - N_{\rm band})$ is the effective number of base terms.

The final remaining piece to mention is our assumption in Equation (5) that the data are centered. This is required so that the simple form of the reference χ_0^2 remains valid. For the multiband model, this assumption requires that the data satisfy Equation (5) *within each band*: equivalently, we could lift this assumption and compute the reference χ_0^2 of the multiband model with an independent floating mean within each band; the results will be identical.

This multiband approach, then, actually comprises a set of models indexed by their value of $N_{\rm base}$ and $N_{\rm band}$. The most fundamental models have $(N_{\rm base}, N_{\rm band}) = (1, 0)$ and $(0, 1)$, which we will call the *shared-phase* and *multi-phase* models, respectively. In the shared-phase model, all variability is assumed to be shared between the bands, with only the fixed offset between them allowed to float. In the multi-phase model, each band has independent variability around a shared fixed offset.

5.1. Relationship of Multiband and Single-band Approaches

With this formalism in place, we can return briefly to the naïve models discussed at the beginning of Section 5. The first, which ignores band information, is simply a standard Lomb–Scargle over the heterogeneous data. The second, in which each band is fit independently, turns out to be equivalent to the $(N_{\rm base}, N_{\rm band}) = (0, 1)$ model defined above. Here the base model is a simple global offset which is degenerate with the offsets in each band, so that the design matrix X_ω can be straightforwardly rearranged as block-diagonal. A block-diagonal design matrix in a linear model indicates that components of the model are being solved independently: here these independent components amount to the single-band floating-mean model from Section 2, fit independently for each of the K bands.

For band k, we will denote the single-band floating-mean periodogram as

$$P_N^{(k)}(\omega) = 1 - \frac{\chi^2_{{\rm min},k}(\omega)}{\chi^2_{0,k}}. \tag{21}$$

The full multiband periodogram is given by

$$P_N^{(0,1)}(\omega) = 1 - \frac{\sum_{k=1}^{K} \chi^2_{\min,k}(\omega)}{\sum_{k=1}^{K} \chi^2_{0,k}} \tag{22}$$

and it can be shown straightforwaredly that $P_N^{(0,1)}$ can be constructed as a weighted sum of $P_N^{(k)}$:

$$P_N^{(0,1)}(\omega) = \frac{\sum_{k=1}^{K} \chi^2_{0,k} P_N^{(k)}}{\sum_{k=1}^{K} \chi^2_{0,k}}. \tag{23}$$

Thus the $(N_{\rm base}, N_{\rm band}) = (0, 1)$ multiband periodogram is identical to a weighted sum of standard periodograms in each band, where the weights $\chi^2_{0,k}$ are a reflection of both the number of measurements in each band and how much those measurements deviate from a simple constant reference model.

5.2. Multiband Periodogram for Simulated Data

Before applying the multiband method to real data, we will explore its effectiveness on a simulated RR Lyrae light curve. The upper panels of Figure 5 show a multiband version of the simulated RR Lyrae light curve from Figure 1. The upper-left panel shows 60 nights of observations spread over a six-month period, and for each night all five bands (*u, g, r, i, z*) are recorded. Using the typical approach from the literature, we individually compute the standard normalized periodogram within each band: the results are shown in the upper-right panel. The data are sampled well enough that a distinct period of 0.622 days can be recognized within each individual band up to the aliasing effect discussed in Section 3.3. Previous studies have made use of the information in multiple bands to choose between aliases and estimate uncertainties in determined periods (e.g., Sesar et al. 2010; Oluseyi et al. 2012). While this approach is sufficient for well-sampled data, it becomes problematic when the multiband data are sparsely sampled.

The lower panels of Figure 5 show the same 60 nights of data, except with only a *single* band observation recorded each night. The lower-left panel shows the observations as a function of phase, and the lower-right panels show the periodograms derived from the data. With only 12 observations for each individual band, it is clear that there is not enough data to accurately determine the period within each single band. The shared-phase $(N_{\rm base}, N_{\rm band}) = (1, 0)$ multiband approach, shown in the lower-right panel, fits a single model to the full data and clearly recovers the true frequency of 0.622 days. The key result is that while methods based on the standard periodogram are suitable for densely sampled data, the multiband periodogram is superior for sparsely sampled multiband observations.

This shared-phase (1, 0) model is only one of the possible multiband options, however: Figure 6 compares multiband fits to this data for models with various choices of $(N_{\rm base}, N_{\rm band})$. We see here many of the characteristics noted above for single-band models: as discussed in Section 4.2, increasing the number of Fourier terms leads to power at multiples of the fundamental period, and increased model complexity (roughly indexed by the effective number of free parameters $M^{\rm eff}$) tends to increase the background level of the periodogram, obscuring significant peaks. For this reason, models with $N_{\rm base} > N_{\rm band}$ are the most promising: they allow a flexible fit with minimal model complexity. Motivated by this, in the next section we will apply the simplest of this class of models, the (1, 0) shared-phase model, to data from Stripe 82 of the Sloan Digital Sky Survey (SDSS).

6. APPLICATION TO STRIPE 82 RR LYRAE

Stripe 82 is a 300 square degree equatorial region of the sky that was repeatedly imaged through multiple bandpasses during phase II of the SDSS (SDSS II; see Sesar et al. 2007). Here we consider the SDSS II observations of 483 RR Lyrae stars compiled and studied by S10, in which periods for these stars were determined based on empirically derived light curve templates. Because the template-fitting method is extremely computationally intensive, S10 first determined candidate periods by taking the top five results of the SuperSmoother (Reimann 1994) algorithm applied to the *g* band; template fits were then performed at each candidate period and the period with the best template fit was reported as the true period. In this section, we make use of this data set to quantitatively evaluate the effectiveness of the multiband periodogram approach.

6.1. Densely Sampled Multiband Data

The full S10 RR Lyrae data set consists of 483 objects with an average of 55 observations in each of the five SDSS *ugriz* bands spread over just under 10 years. In the upper panels of Figure 7 we show the observed data for one of these objects, along with the periodogram derived with the single-band SuperSmoother model[5] and the shared-phase (0, 1) multiband model. Here we have a case that is analogous to that shown for simulated data in the top panels of Figure 5: each band has enough data to easily locate candidate peaks, the best of which is selected via the S10 template-fitting procedure.

The lower panels of Figure 7 compare the S10 period with the best periods obtained from the one-band SuperSmoother (lower-left) and from the shared-phase multiband model (lower-right). To guide the eye, the figure includes indicators of the locations of the beat aliases (dotted lines) and the first harmonic aliases (dashed lines) of the S10 period. Numerical results are summarized in the upper rows of Table 1.

The best-fit SuperSmoother period matches the S10 period in 87% of cases (421/483), while the best-fit multiband period matches the S10 period in 79% of cases (382/483). The modes of failure are instructive: when the SuperSmoother model misses the S10 period, it tends to land on a harmonic alias (i.e., the dashed line). This is due to the flexibility of SuperSmoother: a doubled period spreads the points out, leading to fewer constraints in each neighborhood and thus a smaller average residual around the model. In other words, the SuperSmoother tends to over-fit data that are sparsely sampled. On the other hand, when the multiband model misses the S10 period, it tends to land on a beat alias between the S10 period and the one-day observing cadence (i.e., the dotted lines). This is due to the fact that the single-frequency periodic model is biased, and significantly under-fits the data: it cannot distinguish residuals due to underfitting from residuals due to window function effects.

In both models, the S10 period appears among the top five periods 99% of the time: 477/483 for the SuperSmoother

[5] The SuperSmoother "periodogram" $P_{\rm SS}$ is constructed from the minimum sum of weighted model residuals $\bar{r}_{\rm min}$ analogous with Equation (8): $P_{\rm SS}(\omega) = 1 - \bar{r}_{\rm min}(\omega)/\bar{r}_0$, where $\bar{r}_0$ is the mean absolute residual around a constant model.

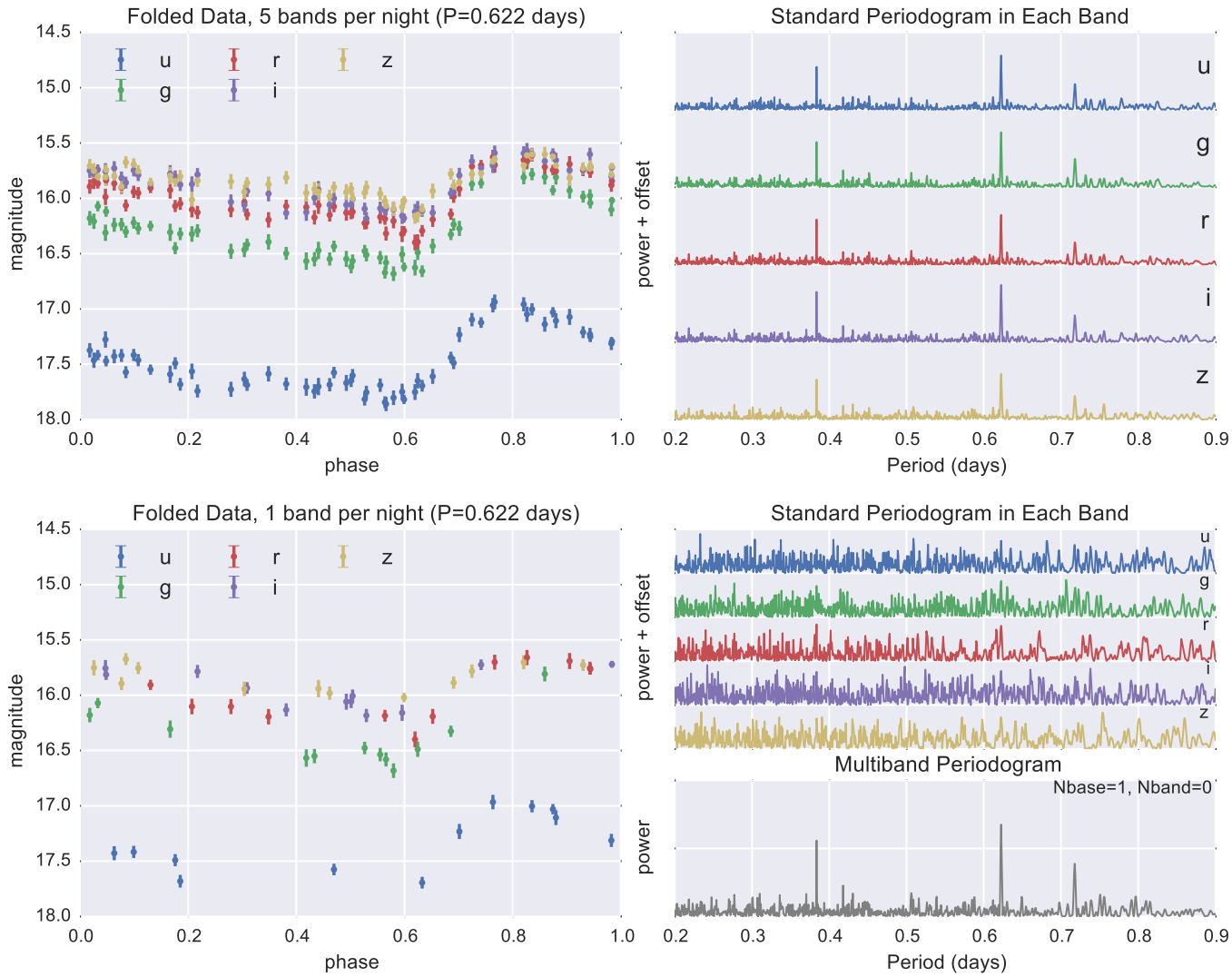

Figure 5. Illustration of the performance of the multiband periodogram. The upper panels show simulated *ugriz* observations of an RR Lyrae light curve in which all five bands are observed each night. With 60 observations in each band, a periodogram computed from any single band is sufficient to determine the true period of 0.622 days. The lower panels show the same data, except with only a single *ugriz* band observed each night (i.e., 12 observations per band). In this case, no single band has enough information to detect the period. The shared-phase multiband approach of Section 5 (lower-right panel) combines the information from all five bands and results in a significant detection of the true period. This indicates that while methods based on the standard periodogram are suitable for densely sampled multiband data, the multiband periodogram is superior for sparsely sampled multiband observations.

periodogram and 480/483 for the multiband periodogram.[6] This suggests that had S10 used the multiband Lomb–Scargle rather than the SuperSmoother in the first pass for their study, the final results presented there would be for the most part unchanged.

The results of this subsection show that the shared-phase multiband approach is comparable to the single-band SuperSmoother approach for densely sampled multiband data, although it has a tendency to get fooled by structure in the survey window. A correction for this phenomenon based on the estimated window power may alleviate this problem (see Roberts et al. 1987 for an example of such an approach) though in practice, selecting from among the top five peaks appears to be sufficient.

[6] We might expect this correspondence to be 100% in the case of the *g*-band SuperSmoother, which was the model used in the first pass of the S10 computation. This discrepancy here is likely due to the slightly different SuperSmoother implementations used in S10 and in this work. Objects showing this discrepancy are those with very low signal-to-noise.

6.2. Sparsely Sampled Multiband Data

Above we saw that the multiband model is comparable to methods from the literature for densely sampled data. Where we expect the multiband approach to gain an advantage is when the data are sparsely sampled, with data through only a single band at each observation time. To simulate this, we reduce the size of the Stripe 82 RR Lyrae data set by a factor of five, keeping only a single band of imaging each night: an average of 11 observations of each object per band. This is much closer to the type of data that will be available in future multiband time-domain surveys.

The upper panels of Figure 8 show an example light curve from this reduced data set along with the SuperSmoother and multiband periodograms derived from this data. Analogously to the lower panels of Figure 5, the single-band SuperSmoother model loses the true period within the noise, while the shared-phase multiband model still shows prominent signal near the S10 period.

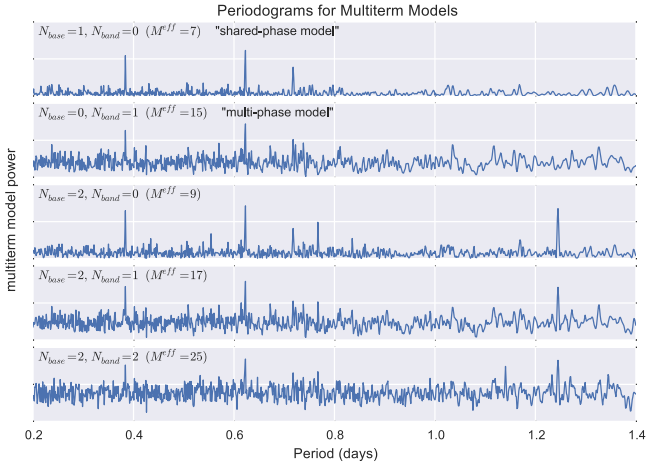

Figure 6. Comparison of the periodograms produced by various multiband models. The data are the same as those used in Figure 5. $N_{\rm base}$ gives the number of Fourier terms in the base model, and $N_{\rm band}$ gives the number of Fourier terms used to fit the residuals around this model within each band. The characteristics discussed with previous figures are also seen here: in particular, the level of "background noise" in the periodogram grows with the model complexity M.

The lower panels of Figure 8 show the relationship between the S10 periods (based on the full data set) and the periods derived with each model from this reduced data set, and these results are summarized in the lower rows of Table 1. It is clear that the SuperSmoother model is simply over-fitting noise with these few data points: the top period matches S10 in only 23% of cases (compared to 87% with the full data set), and the top five periods contain the S10 period only 45% of the time. The failure mode is much less predictable as well: rather than being clustered near aliases, most of the period determinations are scattered seemingly randomly around the parameter space.

While the multiband method performed comparably to the S10 method on dense data, it far outperforms S10 on the sparse data set. Even with an 80% reduction in the number of observations, the multiband method matches the S10 period 64% of the time (compared to 79% with the full data set), and the top five peaks contain the S10 period 94% of the time (compared to 99% with the full data set). This performance is due to the fact that the multiband algorithm has relatively few parameters, yet is able to flexibly accommodate noisy data from multiple observing bands. In particular, this suggests that with the multiterm periodogram, the S10 analysis could have been done effectively with only a small fraction of the available data. This bodes well for future surveys, where data on variable stars will be much more sparsely sampled.

6.3. Potential Improvements to the Multiband Method

A well-known (though often unrecognized) difficulty of Lomb–Scargle-type periodograms on unevenly sampled data is that they do not measure the power of the signal in question but the power of the signal *convolved with observing with the survey window function*. For regularly sampled time series, this convolution is the source of the perfect aliasing beyond the Nyquist sampling limit; for non-regular sampling, this aliasing generally happens to some degree at *all* frequencies! Because of this, even a signal with a single well-defined period will result in a Lomb–Scargle periodogram with multiple maxima at locations that depend on both the underlying signal and the precise observing window.

The multiband periodogram, as a generalization of Lomb–Scargle, shares this difficulty: it tends to respond to frequency structure in the window function as well as frequency structure in the data. This can be viewed as a result of the very model simplicity which causes its success in the case of sparse multiband data: it cannot disentangle bias in the model from bias due to features in the survey window.

This could potentially be accounted for by correcting for the effect of the estimated window function; one potential method for this involves estimating the deconvolution of the window power and the observed power (Roberts et al. 1987). It may also be possible to propose a multiband extension of, e.g., CARMA (Kelly et al. 2014) or another forward-modeling approach to detecting periodicity.

Another potentially fruitful avenue of research that we do not study here is the adjustment of the regularization terms in the model, and the application of other types of regularization to the higher-order periodogram. In particular, L1 regularization (also known as Lasso regression) could lead to interesting results: L1 regularization is similar in spirit to the Tikhonov regularization discussed in Section 4.3, but tends toward sparsity in the model parameters (see, e.g., Ivezić et al. 2014, for a discussion). Such an approach could provide a useful tradeoff between model complexity and bias in the case of higher-order truncated Fourier models, though comes at a higher computational cost.

Another potentially interesting extension of the multiband case would be to define and make use of physically motivated priors in the light-curve shape. This approach could allow the model bias to be decreased without a commensurate increase in model complexity, which is what causes poor performance in the case of sparsely sampled noisy data. As an example of such a physically motivated prior, consider that the paths of RR Lyrae stars through color–color and color–magnitude space are constrained by known astrophysical processes in the structure of the stars (e.g., see Figure 5 in Szabó et al. 2014). Making use of this information could help break degeneracies in period determination with higher-order models.

7. PROSPECTS FOR MULTIBAND PERIODOGRAMS WITH LSST

Previously, Oluseyi et al. (2012) evaluated the prospects of period finding in early LSST data, and found results which were not encouraging. Using the conservative criterion of a 2/3 majority among the top single-band SuperSmoother periods in the g, r, and i bands, they showed that, depending on spectral type, finding reliable periods for the brightest ($g \sim 20$) RR Lyrae stars will require several years of LSST data, while periods for some of the faintest ($g \sim 25$) stars will not be reliable with even 10 years of data!

One potential remedy is to move away from general models like SuperSmoother and Lomb–Scargle to specific template-fitting methods such as those used in S10. Indeed, such methods perform well even for sparsely sampled multiband data such as those from the PanSTARRS survey; the primary drawback is that such blind template fits are computationally extremely expensive: they involve nonlinear optimizations over each of several hundred candidate templates at each of tens of thousands of candidate frequencies (B. Sesar 2015, private communication). Thus the template-fitting method, though it can produce accurate periods, in practice requires several hours of CPU time for a well-sampled period grid for a single source

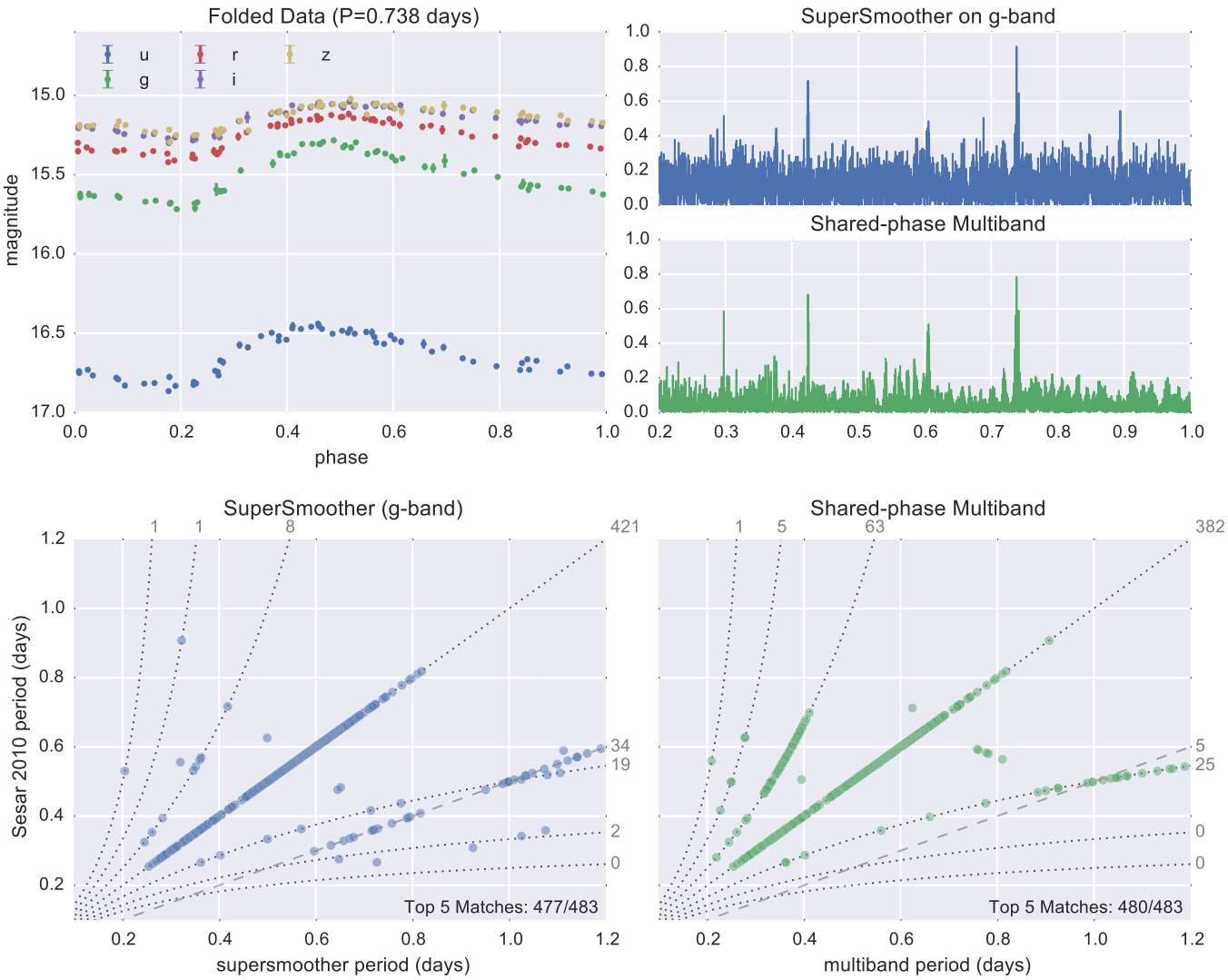

Figure 7. Comparison of the multiband algorithm and single-band SuperSmoother algorithm on 483 well-sampled RR Lyrae light curves from Stripe 82. The upper panels show a representative light curve and periodogram fits, while the bottom panels compare the derived periods to the template-based periods reported in S10. Shown for reference are the beat aliases (dotted lines) and the first harmonic alias (dashed lines): numbers along the top and right edges of the panels indicate the number of points aligned with each trend. The single-band SuperSmoother model tends to err toward harmonic aliases, while the multiband model tends to err toward beat frequency aliases. Both methods find the correct period among the top five significant peaks around 99% of the time. This suggests that for densely sampled multiband surveys, the multiband periodogram will match the results of standard methods (but see Figure 8).

Table 1
Period Determination from Dense and Sparse Data (483 Total)

Data	Method	Match Among Top 5	Top Peak Matches	Beat Aliases	Harmonic Aliases
Dense data (Figure 7)	*g*-band SuperSmoother	477 (98.8%)	421 (87.2%)	31	34
	Multi-band periodogram	480 (99.4%)	382 (79.1%)	94	5
Sparse data (Figure 8)	*g*-band SuperSmoother	219 (45.3%)	113 (23.4%)	101	4
	Multi-band periodogram	449 (93.0%)	308 (63.8%)	136	7

(compared to several seconds for the multiband periodogram proposed here). Note that several hours per object is orders-of-magnitude too slow in the case of LSST; to estimate periods for a billion stars on a 1000 core machine in a year requires a compute-time budget of only 30 s per light curve.

Because of the computational expense of the pure template-fitting method, when working with SDSS II data S10 performed a first-pass with a single-band SuperSmoother to establish candidate periods, which were in turn evaluated with a template-fitting approach. Here we show that such a hybrid strategy of combining the multiband periodogram and the S10 template fits will be useful for determining periodicity of variables in early LSST data releases, greatly improving on the outlook presented in Oluseyi et al. (2012).

We suggest the following procedure for determining periods in future multiband data sets:

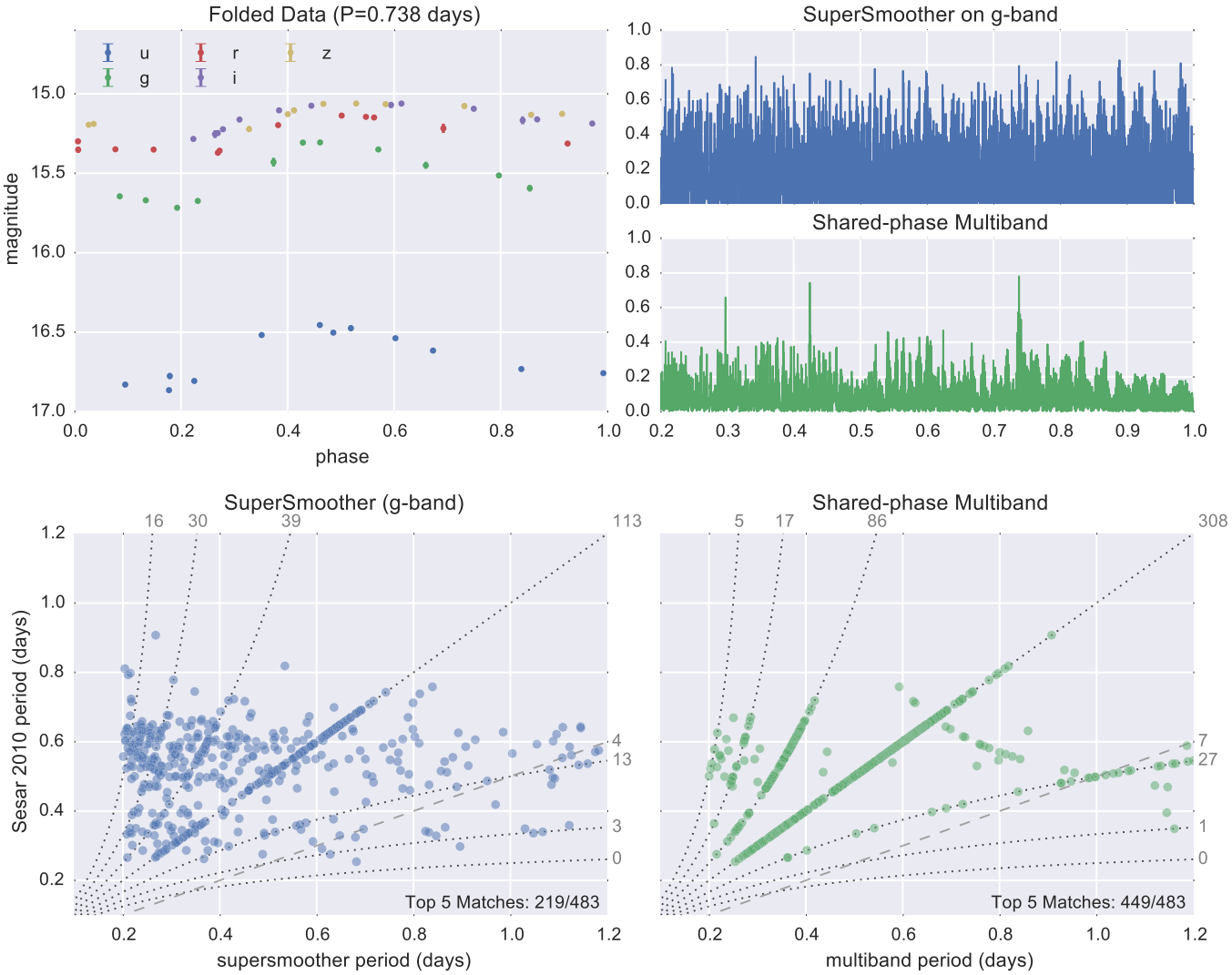

Figure 8. This figure repeats the experiment shown in Figure 7 (see caption there for description), but the data are artificially reduced to only a single-band observation on each evening, a situation reflective of the observing cadence of future large-scale surveys. In this case, the single-band SuperSmoother strategy used as a first pass in S10 fails: there is simply not enough data in each band to recover an accurate period estimate. The correct period is among the top five candidates in fewer than 50% of cases. The shared-phase multiband approach utilizes information from all five bands and returns much more robust results: even with the greatly reduced data, the true period is among the top five candidates in 93% of cases. This suggests that for sparsely sampled multiband survey data (such as that expected from LSST) the multiband periodogram will produce superior results when compared to standard methods—see Figure 9.

1. As a first pass, find a set of candidate frequencies using the multiband periodogram. This is a fast linear optimization that can be straightforwardly parallelized.
2. Within these candidate frequencies, use the more costly template-fitting procedure to choose the optimal period from among the handful of candidates.
3. Compute a goodness-of-fit statistic for the best-fit template to determine whether the fit is suitable; if not, then apply the template-fitting procedure across the full period range.

Here we briefly explore simulated LSST observations of RR Lyrae stars in order to gauge the effectiveness of the first step in this strategy; the effectiveness of the template-fitting step will be explored further in future work. Rather than doing the full analysis including the final template fits, we will focus on the ability of the multiband periodogram to quickly provide suitable candidate periods under the assumption that the S10 template algorithm will then select or reject the optimal period from this set.

7.1. LSST Simulations

We use a simulated LSST cadence (Delgado et al. 2006; Ridgway et al. 2012; Jones et al. 2014) in 25 arbitrarily chosen fields that are representative of the anticipated main survey temporal coverage. We simulate a set of 50 RR Lyrae observations with the S10 templates, with a range of apparent magnitudes between $g = 20$ and $g = 24.5$, corresponding to a bright-to-faint range of LSST main-survey observations, and with expected photometric errors computed using Equations (4)–(6) from Ivezić et al. (2008). Given the capability of template-fitting to choose among candidate periods, we use a more relaxed period-matching criterion than in Oluseyi et al. (2012): when evaluating the single-band SuperSmoother, we require that the true period is among the five periods determined independently in the u, g, r, i, z bands; in the multiband case we require that the true period is among the top five peaks in the multiband periodogram.

Figure 9 shows the fraction of stars where this period matching criterion is met as a function of g-band magnitude

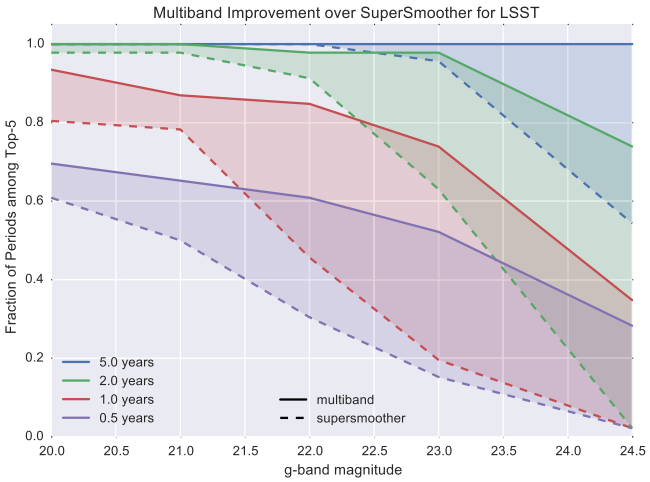

Figure 9. Fraction of periods correctly determined for LSST RR Lyrae as a function of the length of the observing season and the mean g-band magnitude for the multiband periodogram approach (method of this work, solid lines) and single-band SuperSmoother approach (method of Oluseyi et al. 2012, dashed lines). The multiband method is superior to the single-band SuperSmoother approach in all cases and especially for the faintest objects.

and subset of LSST data. The solid lines show the multiband results, the dashed lines show the single-band SuperSmoother results, and the shading helps guide the eye for the sake of comparison. Because of our relaxed matching criteria, even the single-band SuperSmoother results here are much more optimistic than the Oluseyi et al. (2012) results (compare to Figure 15 in that work): the SuperSmoother result here can be considered representative of a best-case scenario for ad hoc single-band fits. Without fail, the multiband result exceeds this best-case single-band result; the improvement is most apparent for faint stars, where the greater model flexibility of the SuperSmoother causes it to over-fit the noisy data.

The performance of the multiband periodogram points to much more promising prospects for science with variable stars than previously reported. In particular, even with only six months of LSST data, we can expect to correctly identify the periods for over 60% of stars brighter than $g = 22$; with the first two years of LSST observations, this increases to nearly 100%; with five years of data, the multiband method identifies the correct period for 100% of even the faintest stars. Part of this improvement is due to the performance of the shared-phase multiband model with noisy data, and part of this improvement is due to the relaxed period-matching constraints enabled by the hybrid approach of periodogram-based and template-based period determination.

8. DISCUSSION AND CONCLUSION

We have derived a multiband version of the classic Lomb–Scargle method for detecting periodicity in an astronomical time series. Experiments on several hundred RR Lyrae stars from the SDSS Stripe 82 data set indicate that this method outperforms methods used previously in the literature, especially for sparsely sampled light curves with only single bands observed each night. While there are potential areas of improvement involving corrections to window function artifacts and accounting for physically motivated priors, the straightforward multiband model outperforms previous ad hoc approaches to multiband data.

Looking forward to future variable star catalogs from PanSTARRS, DES, and LSST, there are two important constraints that any analysis method must meet: the methods must be able to cope with heterogeneous and noisy observations through multiple bandpasses, and the methods must be fast enough to be computable on millions or even billions of objects. The multiband method, through its combination of flexibility and model simplicity, meets the first constraint: as shown above, in the case of sparsely sampled noisy multiband data, it outperforms previous approaches to period determination. It also meets the second constraint: it requires the solution of a simple linear model at each frequency, compared to a rank-based sliding-window model in the case of SuperSmoother, a nonlinear optimization in the case of template-fitting, and a Markov Chain Monte Carlo analysis in the case of CARMA models. In our own benchmarks, we found the multiband method to be several times faster than the single-band SuperSmoother approach and several orders of magnitude faster than the template-fitting approach.

The strengths and weaknesses of the multiband method suggest a hybrid approach to finding periodicity in sparsely sampled multiband data: a first pass with the fast multiband method, followed by a second pass using the more computationally intensive template-fitting method to select among these candidate periods. Despite pessimism in previous studies, our experiments with simulated LSST data indicate that such a hybrid approach will successfully identify periods in the majority of RR Lyrae stars brighter than $g \sim 22.5$ in the first months of the survey, and the majority of the faintest detected stars with several years of data. This finding suggests that the multiband periodogram could have an important role to play in the analysis of variable stars in future multiband surveys.

We have released a Python implementation of the multiband periodogram on GitHub, along with Python code to reproduce all results and figures in this work; this is described in Appendix A. As we were finalizing this manuscript, we were made aware of a preprint of an independent exploration of a similar approach to multiband light curves (Long et al. 2014); we discuss the similarities and differences between these two approaches in Appendix B.

J. T. V. is supported by the University of Washington eScience institute, including grants from the Alfred P. Sloan Foundation, the Gordon and Betty Moore Foundation, and the Washington Research Foundation. The authors thank GitHub for providing free academic accounts that were essential in the development of this work.

APPENDIX A
PYTHON IMPLEMENTATION OF MULTIBAND PERIODOGRAM

The algorithm outlined in this paper is available in `gatspy`, an open-source Python package for general astronomical time-series analysis[7] (Vanderplas 2015a). Along with the periodogram implementation, it also contains code to download all the data used in this work. Code to reproduce this paper, including all figures, is available in a separate repository.[8]

`gatspy` is a pure Python package written to be compatible with both Python 2 and Python 3 and performs fast numerical

[7] http://github.com/astroml/gatspy/
[8] http://github.com/jakevdp/multiband_LS/

computation through dependencies on `numpy` (van der Walt et al. 2011)[9] and `astroML` (Vanderplas et al. 2012),[10] which offer optimized implementations of numerical methods in Python.

The API for the module is largely influenced by that of the `scikit-learn` package (Pedregosa et al. 2011; Buitinck et al. 2013),[11] in which models are Python class objects which can be fit to data with the `fit()` method. Here is a basic example of how you can use `multiband_LS` to download the data used in this paper, fit a multiband model to the data, and compute the power at a few periods:

```
from gatspy.periodic import LombScargleMultiband
import numpy as np
# Fetch the Sesar 2010 RR Lyrae data
from gatspy.datasets import fetch_rrlyrae
data = fetch_rrlyrae()
t, mag, dmag, filts = data.get_lightcurve(data.ids[0])
# Construct the multiband model
model = LombScargleMultiband(Nterms_base = 0,
  Nterms_band = 1)
model.fit(t, mag, dmag, filts)
# Compute power at the following periods
periods = np.linspace(0.2, 1.4, 1000) # periods in days
power = model.periodogram(periods)
```

Other models are available as well. For example, here is how you can compute the periodogram under the SuperSmoother model; this implementation of the SuperSmoother periodogram makes use of the `supersmoother` Python package (Vanderplas 2015b).

```
from gatspy.periodic import supersmoother
# Construct the supersmoother model
model = supersmoother()
gband = (filts == 'g')
model.fit(t[gband], mag[gband], dmag[gband])
# Compute power at the given periods
power = model.periodogram(periods)
```

The models in the `gatspy` package contain many more methods, and much more functionality that what is shown here. For updates, more examples, and more information, visit http://github.com/astroml/gatspy/.

APPENDIX B
COMPARISON WITH LONG (2014)

As we were finishing this study, we learned that another group had released a preprint independently addressing the multiband periodogram case and had come up with a solution very similar to the one presented here (Long et al. 2014, hereafter LCB14). They present two methods, the "Multiband Generalized Lomb–Scargle" (MGLS) which is effectively identical to the (1, 0) multi-phase model here, and the "Penalized Generalized Lomb–Scargle" (PGLS), which is similar in spirit to our (0, 1) shared-phase model.

In the PGLS model, they start with a multi-phase model, fitting independent $N = 1$ term fits to each band, and they apply a nonlinear regularization term that penalizes differences in the amplitude and phase. In terms of the formalism used in this work, the PGLS model minimizes a regularized χ^2 of the form

$$\chi^2_{\rm PGLS} = \sum_{k=1}^{K} \left[\chi^2_{\rm GLS}(D^{(k)}) + J_A(A^{(k)}) + J_\phi(\phi^{(k)}) \right]. \quad (24)$$

where K is the number of bands, $\chi^2_{\rm GLS}(D^{(k)})$ is the χ^2 of the standard floating mean model on the single-band data D^k, and J_A and J_ϕ are regularization/penalty terms that are a function of the amplitude A^k and phase ϕ^k of each model. In terms of our linear model parameters θ^k, this amplitude and phase can be expressed as:

$$A^{(k)} = \sqrt{(\theta_1^{(k)})^2 + (\theta_2^{(k)})^2}$$
$$\phi^{(k)} = \arctan(\theta_2^{(k)}/\theta_1^{(k)}). \quad (25)$$

The selected form of these regularization terms penalizes deviations of the amplitude and phase from a common mean between the bands; in this sense the PGLS model can be considered a conceptual mid-point between our shared-phase and multi-phase models. Within the formalism proposed in the current work, such a mid-point may be alternatively attained by suitably increasing the regularization parameter λ used in our shared-phase model, though the precise nature of the resulting regularization will differ.

Computationally, the PGLS model requires a nonlinear optimization at each frequency ω and is thus much more expensive than the straightforward linear optimization of our shared-phase model. For this reason, LCB14 proposes a clever method by which nested models are used to reduce the number of nonlinear optimizations used: essentially, by showing that the (linear) MGLS χ^2 is a lower-bound of the (nonlinear) PGLS χ^2, it is possible to iteratively reduce the number of PGLS computations required to minimize the χ^2 among a grid of frequencies. Such an optimization could also be applied in the case of our shared-phase model, but is not necessary here due to its already high speed. Nevertheless, when applying the method to a very large number of light curves as in, e.g., LSST, such a computational trick may prove very useful.

Given these important distinctions between the models proposed here and in LCB14, in future work we plan to do a detailed comparison of the two approaches to multiband model regularization.

REFERENCES

Brandt, S. 1970, Statistical and Computational Methods in Data Analysis (Amsterdam: North-Holland)
Bretthorst, G. 1988, Bayesian Spectrum Analysis and Parameter Estimation (Berlin: Springer)
Buitinck, L., Louppe, G., Blondel, M., et al. 2013, arXiv:1309.0238
Cumming, A., Marcy, G. W., & Butler, R. P. 1999, ApJ, 526, 890
Delgado, F., Cook, K., Miller, M., Allsman, R., & Pierfederici, F. 2006, Proc. SPIE, 6270, 62701D
Eyer, L., & Mowlavi, N. 2008, JPhCS, 118, 012010
Ferraz-Mello, S. 1981, AJ, 86, 619
Flaugher, B. 2008, in A Decade of Dark Energy: Spring Symp., ed. N. Pirzkal & H. Ferguson, http://www.stsci.edu/institute/conference/spring2008
Graham, M. J., Drake, A. J., Djorgovski, S. G., et al. 2013, MNRAS, 434, 3423
Hoerl, A. E., & Kennard, R. W. 1970, Technometrics, 12, 55
Ivezić, Ž, Connolly, A., VanderPlas, J., & Gray, A. 2014, Statistics, Data Mining, and Machine Learning in Astronomy: A Practical Python Guide for the Analysis of Survey Data (Princeton, NJ: Princeton Univ. Press)
Ivezić, Ž, Tyson, J. A., Abel, B., et al. 2008, arXiv:0805.2366

[9] http://www.numpy.org
[10] http://www.astroml.org
[11] http://scikit-learn.org

Jaynes, E. 1987, Fundamental Theories of Physics, Vol. 21, ed. C. Smith & G. Erickson (Netherlands: Springer)
Jones, R. L., Yoachim, P., Chandrasekharan, S., et al. 2014, Proc. SPIE, 9149, 91490B
Kaiser, N., Burgett, W., Chambers, K., et al. 2010, Proc. SPIE, 7733, 77330E
Kelly, B. C., Becker, A. C., Sobolewska, M., Siemiginowska, A., & Uttley, P. 2014, ApJ, 788, 33
Lomb, N. R. 1976, Ap&SS, 39, 447
Long, J. P., Chi, E. C., & Baraniuk, R. G. 2014, arXiv:1412.6520
Oluseyi, H. M., Becker, A. C., Culliton, C., et al. 2012, AJ, 144, 9
Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825
Press, W. H., & Rybicki, G. B. 1989, ApJ, 338, 277
Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, Numerical Recipes: The Art of Scientific Computing (3rd ed.; New York: Cambridge Univ. Press)
Reimann, J. D. 1994, PhD thesis, Univ. California, Berkeley
Ridgway, S. T., Chandrasekharan, S., Cook, K. H., et al. 2012, Proc. SPIE, 8448, 844810
Roberts, D. H., Lehar, J., & Dreher, J. W. 1987, AJ, 93, 968
Scargle, J. D. 1982, ApJ, 263, 835
Schuster, A. 1898, TeMag, 3, 13
Sesar, B., Ivezić, Ž, Grammer, S. H., et al. 2010, ApJ, 708, 717
Sesar, B., Ivezić, Ž, Lupton, R. H., et al. 2007, AJ, 134, 2236
Sesar, B., Stuart, J. S., Ivezić, Ž, et al. 2011, AJ, 142, 190
Stellingwerf, R. F. 1978, ApJ, 224, 953
Süveges, M., Sesar, B., Váradi, M., et al. 2012, MNRAS, 424, 2528
Szabó, R., Ivezić, Ž, Kiss, L. L., et al. 2014, ApJ, 780, 92
Tikhonov, A. 1963, Soviet Math. Dokl., 5, 1035
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Vanderplas, J. 2015a, gatspy: General Tools for Astronomical Time Series in Python, Zenodo, doi:10.5281/zenodo.14833
Vanderplas, J. 2015b, SuperSmoother: Efficient Python Implementation of Friedman's SuperSmoother, Zenodo, doi:10.5281/zenodo.14475
Vanderplas, J., Connolly, A., Ivezć, Ž, & Gray, A. 2012, in Conference on Intelligent Data Understanding (CIDU) (Boulder, CO: IEEE), 47
Welch, D. L., & Stetson, P. B. 1993, AJ, 105, 1813
Zechmeister, M., & Kürster, M. 2009, A&A, 496, 577