Reconstruction of Gaussian and log-normal fields with spectral smoothness

Niels Oppermann,* Marco Selig, Michael R. Bell, and Torsten A. Enßlin

Max Planck Institute for Astrophysics, Karl-Schwarzschild-Str. 1, 85741 Garching, Germany

(Dated: October 26, 2012)

We develop a method to infer log-normal random fields from measurement data affected by Gaussian noise. The log-normal model is well suited to describe strictly positive signals with fluctuations whose amplitude varies over several orders of magnitude. We use the formalism of minimum Gibbs free energy to derive an algorithm that uses the signal's correlation structure to regularize the reconstruction. The correlation structure, described by the signal's power spectrum, is thereby reconstructed from the same data set. We further introduce a prior for the power spectrum that enforces spectral smoothness. The appropriateness of this prior in different scenarios is discussed and its effects on the reconstruction's results are demonstrated. We validate the performance of our reconstruction algorithm in a series of one- and two-dimensional test cases with varying degrees of non-linearity and different noise levels.

I. INTRODUCTION

Reconstructing continuous fields from a finite and noisy data set is a problem encountered often in all branches of physics and the geo-sciences. In this paper we develop a method to reconstruct a log-normal field, i.e. a field whose logarithm can be modeled as a Gaussian random field, defined on an arbitrary manifold. We show simple one-dimensional examples as well as planar and spherical two-dimensional cases.

The log-normal model is well suited to describe many physical fields. Its main features are that the field values are guaranteed to be positive, that they may vary across many orders of magnitude, and that spatial correlations may exist. One prominent example from the discipline of astrophysics that exhibits these three features and is therefore often modeled as a log-normal field is the matter density in today's universe [see e.g. 5]. The simplest way to observationally estimate the matter density is to count the galaxies per volume. The relationship between the matter density and the galaxy number counts can be approximated as a Poisson process [e.g. 12, 14], thus making the data likelihood Poissonian as well with the expected number of galaxies per volume element given as a function of the underlying log-normal density field. In the statistical literature, such a combination of a log-normal field with Poissonian statistics is known as a log Gaussian Cox process [17]. This model has been applied in fields as diverse as finance [1], agriculture [3], atmospheric studies [22], and epidemiology [2, 4].

The main motivation for our work, however, comes from astronomical observations of radiative processes. The intensity of radiation coming from different directions in such observations can vary over many orders of magnitude while being spatially correlated. Thus applying a log-normal model for its statistical description

seems natural. While photon counts bring with them a Poissonian likelihood, the observational uncertainty can be approximatively regarded to be Gaussian in the limit of large photon numbers. Therefore, we restrict ourselves to cases in which the measurement noise can be assumed to be additive and Gaussian.

Apart from the noise contribution, we will assume a deterministic linear relationship between the log-normal field and the data. This model is general enough to accomodate a large variety of observational settings, such as targeted point-wise observations across limited areas, convolutions of the log-normal field, interferometric observations leading to a Fourier transformed version of the log-normal field, or any combination of these effects.

The inclusion of spatial correlations in the field model is necessary for an accurate statistical description and will aid in the reconstruction. If the spatial correlation structure is known, knowledge of the field values at some locations can be used to extrapolate the field into regions where the field value has not been measured or the observations have a higher uncertainty. However, in general it is not known a priori how strongly the field is correlated across a given distance. In the case of a statistically homogeneous field, the correlation structure is described by the field's power spectrum. So if one wants to make use of the spatial correlations during the reconstruction, one needs to simultaneously infer the power spectrum. Several techniques have been developed to solve this problem for Gaussian fields [e.g 13, 26]. One such technique was derived within the framework of *information field theory* [7, 16] in [6] and later rederived in [8], where the formalism of minimum Gibbs free energy was introduced and employed to tackle this problem. In the same paper, the problem of inferring a log-normal field with unknown power spectrum in the presence of Poissonian noise was briefly discussed.

Here, we use the formalism of minimum Gibbs free energy to derive filter equations for the estimation of a log-normal field from data including a Gaussian noise

---

* niels@mpa-garching.mpg.de

component. The filter we present results in an estimate of the field, its power spectrum, and the respective uncertainties.

A problem that often arises when reconstructing power spectra is that the reconstructed power spectra do not turn out smooth, especially on scales with little redundancy, despite the true underlying spectra being smooth. One reason why this is not desirable is that sharp kinks and drop-offs in the power-spectrum can lead to ringing effects in the reconstructed field estimate. Another reason is that many of the fields occuring in nature are actually expected to have smoothly varying power over different scales since neighboring scales interact. One prominent example is the power spectrum of a viscous turbulent fluid [15]. Smoothness of the power spectrum is often enforced by an ad-hoc smoothing step in the power spectrum reconstruction [18–20]. Here, we follow an idea presented in [6] and show how to enforce spectral smoothness by introducing an appropriate prior for the power spectrum that punishes non-smooth spectra. We demonstrate the feasibility of this approach using one specific example for such a smoothness prior both for the reconstruction of Gaussian and log-normal fields and discuss the range of applicability of the chosen spectral smoothness prior. Our approach also allows for the estimation of a complete uncertainty matrix for the estimated power spectrum.

We first develop the formalism of the spectral smoothness prior for the case of Gaussian signal fields in Sec. II, where we show how to derive the filter formulas of [6] and [8] in a different way that easily accomodates an additional smoothness prior. After demonstrating the workings of the spectral smoothness prior, we use the Gibbs free energy formalism to derive filter formulas for the log-normal case and transfer the spectral smoothness results to this case in Sec. III. We demonstrate the performance of our log-normal reconstruction algorithm in a variety of test cases in Sec. III B and conclude in Sec. IV.

II. RECONSTRUCTING GAUSSIAN FIELDS WITH SPECTRAL SMOOTHNESS

First, we need to introduce some basic assumptions and notation. We mainly follow the notation that is used throughout the literature on *information field theory* [e.g. 7].

Throughout the paper, we assume that we are analyzing a set of data $d = (d_1, \ldots, d_r) \in \mathbb{R}^r$ that depends on an underlying signal field $s : \mathcal{M} \longrightarrow \mathbb{R}$, subject to additive noise $n = (n_1, \ldots, n_r) \in \mathbb{R}^r$,

$$d = f(s) + n. \tag{1}$$

Here, $\mathcal{M}$ is the discrete or continuous space or manifold that the signal is defined on. In Sec. III we will discuss the examples $\mathcal{M} = \mathcal{S}^1$, $\mathcal{S}^2$, and $\mathcal{T}^2$. In this section, we assume the relationship between signal field and data to

be linear, described by a response operator R, so that

$$d = Rs + n. \tag{2}$$

In most applications, the response operator will include some instrumental or observational effects such as observations in specific locations of $\mathcal{M}$, convolutions of s with an instrumental response function, or a Fourier transform of the signal field. The only restriction that we make here is that the operation that generates the data from the signal has to be linear in the signal. Finally, we restrict ourselves to real-valued signals only in the interest of notational simplicity. All our results can be straightforwardly generalized to complex-valued signal fields.

For the noise n, we assume Gaussian statistics,

$$n \hookleftarrow \mathcal{G}(n, N), \tag{3}$$

described by a not necessarily diagonal covariance matrix N. Here, $\mathcal{G}(\phi, \Phi)$ denotes a multi-variate Gaussian probability distribution function with covariance Φ,

$$\mathcal{G}(\phi, \Phi) = \frac{1}{|2\pi\Phi|^{1/2}} \exp\left(-\frac{1}{2}\phi^\dagger \Phi^{-1} \phi\right). \tag{4}$$

We use the $\dagger$-symbol to denote a transposed (and in general complex conjugated) quantity and take the limit of an infinite-dimensional Gaussian whenever the argument is a continuous field, with the appropriate limit of the scalar product

$$\phi^\dagger \psi = \int_{\mathcal{M}} dx \, \overline{\phi(x)} \, \psi(x) \quad \forall \, \phi, \psi : \, \mathcal{M} \longrightarrow \mathbb{R}. \tag{5}$$

In this section we will deal with the case that the signal field s can also be regarded – or at least approximated – as a zero-mean Gaussian random field with covariance S,

$$s \hookleftarrow \mathcal{G}(s, S). \tag{6}$$

It can be straightforwardly shown [e.g. 7] that the posterior mean m of the signal field under these assumptions is given by

$$m = \int \mathcal{D}s \, s \, \mathcal{P}(s|d) = Dj. \tag{7}$$

Here, $\int \mathcal{D}s$ denotes an integral over the configuration space of all possible signal realizations. The operator

$$D = \left(S^{-1} + R^\dagger N^{-1} R\right)^{-1} \tag{8}$$

is known as the *information propagator* and the field

$$j = R^\dagger N^{-1} d \tag{9}$$

is called *information source*.

In these formulas, the presence of the Gaussian signal prior described by the signal covariance serves as a means of regularization of the sought continuous signal

field reconstruction which otherwise is under-determined when only constrained by the finite and noisy data set. However, in most physical applications, the signal covariance is not known a priori. The problem of reconstructing Gaussian fields with unknown signal covariances has in principle been solved [6, 8, 13, 26], and even an unknown noise covariance can be overcome [21].

Enßlin & Weig [8] use the formalism of minimum Gibbs free energy to derive filter formulas to be applied to the data set when the signal's covariance is unknown. We will recapture this formalism briefly in Sec. III, where we employ it to reconstruct log-normal signal fields.

Under the assumption of statistical homogeneity and isotropy, the unknown signal covariance becomes diagonal in the harmonic eigenbasis, i.e. the Fourier basis for signals defined on Euclidean space and the spherical harmonics basis for signals defined on the sphere. In the following, we denote as $\vec{k}$ the vector of parameters determining one mode in the harmonic decomposition, i.e. $\vec{k} = (k_1, \ldots, k_n)$ for n-dimensional Euclidean space and $\vec{k} = (\ell, m)$ for the two-sphere, where ℓ is the angular momentum quantum number and m the azimuthal one. Furthermore, k shall stand for the scale of the harmonic component, i.e. $k = \sqrt{k_1^2 + \cdots + k_n^2}$ and $k = \ell$ for $\mathbb{R}^n$ and $\mathcal{S}^2$, respectively.

Due to the isotropy assumption, the diagonal of the signal covariance matrix, the signal's power spectrum P_k, depends only on the scale k and not on the specific mode $\vec{k}$, i.e.

$$S_{\vec{k}\vec{k}'} = \delta_{\vec{k}\vec{k}'} P_k. \tag{10}$$

Under these symmetry assumptions, the filter formulas derived by Enßlin & Weig [8] can be written as

$$m = Dj, \tag{11}$$

$$P_k = \frac{1}{\alpha_k - 1 + \frac{\rho_k}{2}} \left(q_k + \frac{1}{2} \sum_{\{\vec{k}'|k'=k\}} \left(|m_{\vec{k}'}|^2 + D_{\vec{k}'\vec{k}'} \right) \right). \tag{12}$$

Note that the first of these equations is the same as in the case of a known signal covariance, Eq. (7), a generalized Wiener filter [7]. Here, α_k and q_k are parameters used to determine the priors for the spectral coefficients P_k (see Sec. III for details) and

$$\rho_k = \sum_{\{\vec{k}'|k'=k\}} 1 \tag{13}$$

is the number of degrees of freedom on the scale k.

This formalism has been successfully applied for astrophysical reconstructions [19, 20] and in a more abstract computational setting [24]. In these applications, Eqs. (11) and (12) were simply iterated. However, they were supplemented with an additional ad-hoc smoothing step for the power spectrum as part of each iteration.
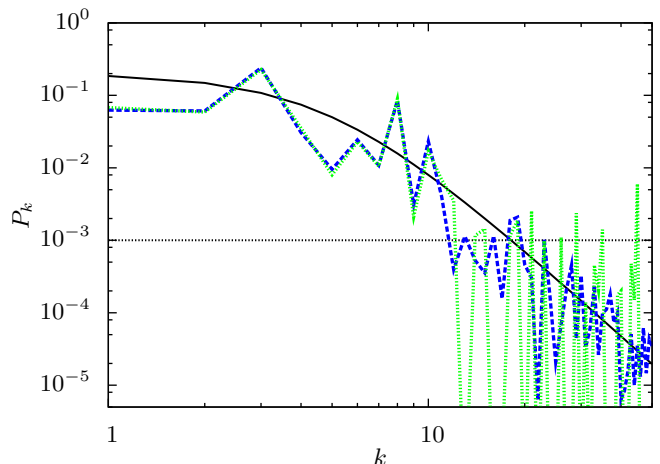

FIG. 1. Power spectra of the one-dimensional test-case. The black solid line shows the theoretical power spectrum, the blue dashed line the power in the randomly drawn signal realization studied here, and the green dotted line shows the power spectrum reconstructed without smoothing. The horizontal dotted line indicates the noise power, given by $\sigma_n^2/100$.

To illustrate the usefulness of a smoothing step, we set up a simple one-dimensional test-case. Here, we assume that our signal field is defined on the one-sphere $\mathcal{S}^1$, i.e. the interval $[0, 1)$ with periodic boundary conditions, which we discretize into 100 pixels. We make up a power spectrum for the signal of the form

$$P_k = P_0 \left(1 + \left(\frac{k}{k_0} \right)^2 \right)^{-\gamma/2}, \tag{14}$$

where we choose $P_0 = 0.2$, $k_0 = 5$, and $\gamma = 4$. Figures 1 and 2 show the power spectrum and the signal realization randomly drawn from it, respectively.

Furthermore, we assume that we have measured the signal field in each of the 100 pixels once, subject to homogeneous and uncorrelated Gaussian noise with variance $\sigma_n^2 = 0.1$. In the formalism of Eq. (2), this corresponds to $R = \mathbb{1}$ and $N = \sigma_n^2 \mathbb{1}$. The data realization is also shown in Fig. 2.

We overplot in Figs. 1 and 2 the power spectrum and map reconstruction, respectively, obtained by iterating Eqs. (11) and (12) without smoothing. While the signal reconstruction suffers from some too pronounced small-scale fluctuations, the reconstructed power spectrum fluctuates wildly for $k > 10$ and is therefore not trustworthy at all on these scales. The scales on which the reconstructed power by chance peaks above the noise level are the ones that are especially misrepresented in the reconstructed signal field.

This example should serve to illustrate that some kind of spectral smoothing is necessary, especially in cases where one is interested in the power spectrum itself. In Fig. 2 we also show a comparison to the Wiener filter reconstruction, i.e. Eq. (7), using the true power spectrum. In the bottom panel of Fig. 2 it can be seen that the residuals, i.e. the differences between the true signal
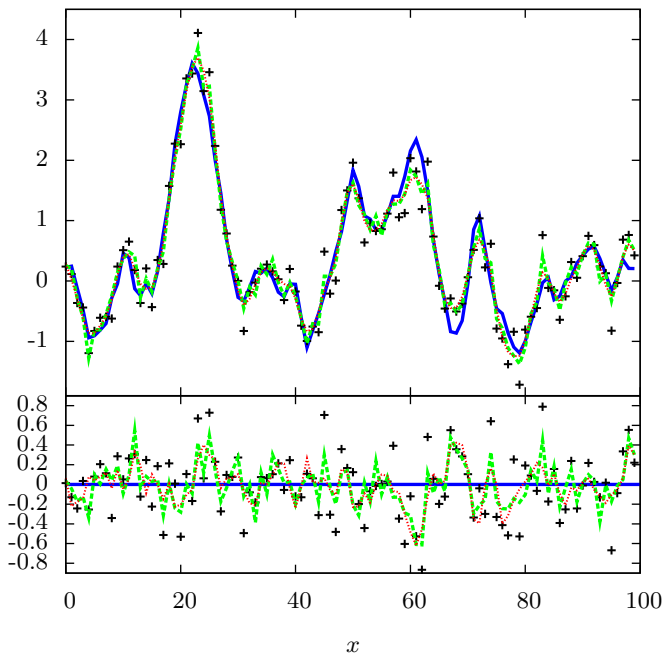
FIG. 2. Signal reconstruction in the one-dimensional test-case. The blue solid line shows the randomly drawn signal realization, the crosses are noisy data points used to do the reconstruction, the green dashed line shows the signal reconstruction obtained without spectral smoothing and the red dotted line is the Wiener filter reconstruction. The lower panel shows the same with the true signal substracted.

and its reconstruction, are reduced if a power spectrum is used that is closer to the true one. Smoothness is one aspect of the true power spectrum that can be used to constrain its reconstruction.

The two-point correlation function of the signal provides another way of looking at the smoothness property of its power spectrum. For a statistically homogeneous signal, the power spectrum is the Fourier transform of the correlation function and vice versa. Thus, a power spectrum that exhibits fluctuations on arbitrarily small scales in k-space corresponds to a signal that exhibits correlations over arbitrarily large scales in position space. Turning this argument around, any signal with correlations only over a finite range in position space or at least correlations that are decreasing rapidly with distance will have a power spectrum that does not exhibit any features on arbitrarily small scales in k-space. Thus, the power spectrum can be expected to be smooth on the reciprocal scale of the typical correlation length scale.

In the remainder of this section, we will show how to incorporate a prior enforcing spectral smoothness into the formalism presented thus far.

A. Reconstruction as a combination of *posterior mean* and *maximum a posteriori*

Before incorporating a spectral smoothness prior into the derivation of Eqs. (11) and (12), we will rederive them in a way different from the one presented in [6, 8].

As was already mentioned, Eq. (11) corresponds to the posterior mean of the signal under the assumption of a known covariance matrix, i.e. a known power spectrum. Since the posterior probability distribution is Gaussian in this case, its mean also maximizes the probability, so that m is the posterior mean and the maximum a posteriori solution at the same time,

$$m = \int \mathcal{D}s \, s \, \mathcal{P}(s|d, P) = \underset{s}{\operatorname{argmax}} \left\{ \mathcal{P}(s|d, P) \right\}. \quad (15)$$

Here, we will show that Eq. (12) can be written as a maximum a posteriori solution as well, considering the signal-marginalized posterior

$$\mathcal{P}(P|d) = \int \mathcal{D}s \, \mathcal{P}(s, P|d). \quad (16)$$

In order to do this, we first have to define a prior for the power spectrum P. In accordance with [6, 8], we choose independent inverse-gamma distributions for each spectral component P_k,

$$\mathcal{P}(P) = \prod_k \mathcal{P}_{\mathrm{IG}}(P_k)$$

$$= \prod_k \frac{1}{q_k \Gamma(\alpha_k - 1)} \left(\frac{P_k}{q_k} \right)^{-\alpha_k} \exp\left(-\frac{q_k}{P_k} \right). \quad (17)$$

Here, $\Gamma(\cdot)$ denotes the gamma function, q_k is a parameter defining the location of an exponential cut-off at low values of P_k, and α_k defines the slope of a power-law for large values of P_k. By tuning these parameters, the prior can be narrowed or widened according to the a priori knowledge about the power spectrum. Taking the limit $q_k \to 0$ and $\alpha_k \to 1$ turns the inverse-gamma distribution into Jeffreys prior which is flat on a logarithmic scale. In the examples presented in this paper we always take this limit in the final filter formulas.

In the following, we will work with the logarithmic spectral components

$$p_k = \log P_k. \quad (18)$$

The corresponding prior for these can be straightforwardly derived from the conservation of probability under variable transformations and reads

$$\mathcal{P}(p) = \mathcal{P}(P) \left| \frac{\mathrm{d}P}{\mathrm{d}p} \right|$$

$$= \prod_k \frac{q_k^{\alpha_k - 1}}{\Gamma(\alpha_k - 1)} e^{-(\alpha_k - 1)p_k} \exp\left(-q_k e^{-p_k} \right). \quad (19)$$

Using this prior, we can calculate the signal-marginalized probability of data and power spectrum, $\mathcal{P}(d, p)$, and its negative logarithm, the Hamiltonian

$$
\begin{aligned}
H(d, p) &= -\log \mathcal{P}(d, p) \\
&= -\log \int \mathcal{D}s \, \mathcal{G}(d - Rs, N) \, \mathcal{G}(s, S) \, \mathcal{P}(p) \\
&= \frac{1}{2} \mathrm{tr} \, (\log S) - \frac{1}{2} \mathrm{tr} \, (\log D) - \frac{1}{2} j^\dagger D j \\
&\quad + \sum_k \left((\alpha_k - 1) \, p_k + q_k \mathrm{e}^{-p_k} \right) + \mathrm{const.,} \quad (20)
\end{aligned}
$$

where we have made use of the definitions (8) and (9) and have collected all terms that do not depend on the power spectrum into an unimportant additive constant. Using the spectral dependence of the signal covariance matrix in a statistically homogeneous and isotropic setting, Eq. (10), we can take the derivative of the Hamiltonian with respect to one p_k and equate it to zero, thus maximizing the posterior probability of the logarithmic power spectrum. The resulting equation is exactly Eq. (12) for $P_k = \mathrm{e}^{p_k}$ if one makes the identification $m = Dj$.

Thus we have shown that the filter formulas, Eqs. (11) and (12), can be derived as a combination of posterior mean for the signal reconstruction and maximum a posteriori for the power spectrum. This effectively means that we have made the approximation

$$
\begin{aligned}
m &= \int \mathcal{D}s \, s \, \mathcal{P}(s|d) \\
&= \int \mathcal{D}s \int \mathcal{D}p \, s \, \mathcal{P}(s|d, p) \mathcal{P}(p|d) \\
&\approx \int \mathcal{D}s \int \mathcal{D}p \, s \, \mathcal{P}(s|d, p) \, \delta \left(p - p^{(\mathrm{MAP})} \right), \quad (21)
\end{aligned}
$$

i.e. we have approximated the posterior probability distribution for the power spectrum with a delta distribution centered on its maximum.

It may be worth noting that in the formalism of the maximum a posteriori solution for the power spectrum, it is straightforward to derive a rough uncertainty estimate as well. The Hessian of the Hamiltonian gives the curvature of the posterior probability distribution and its inverse can thus be regarded as an uncertainty matrix. For the Hamiltonian given in Eq. (20) we obtain

$$
\begin{aligned}
\left. \frac{\partial^2 H(d, p)}{\partial p_k \partial p_{k'}} \right|_{p=p^{(\mathrm{MAP})}} &= \left(\alpha_k - 1 + \frac{\rho_k}{2} \right) \delta_{kk'} \\
&\quad - \frac{1}{2} \mathrm{e}^{-p_k - p_{k'}} \sum_{\substack{\{\vec{q}|q=k\} \\ \{\vec{q}'|q'=k'\}}} \left(2 \, \Re \big(m_{\vec{q}} m_{\vec{q}'}^* D_{\vec{q}\vec{q}'} \big) + |D_{\vec{q}\vec{q}'}|^2 \right) \bigg|_{p=p^{(\mathrm{MAP})}}
\end{aligned}
$$

$$
(22)
$$

where $\Re(\cdot)$ denotes the real part of a complex number.

B. Spectral smoothness priors

Here, we show how to incorporate a spectral smoothness prior into the formalism developed in the previous section. We do this by augmenting the inverse-gamma distributions previously assumed as the spectral prior with a probability distribution that enforces smoothness of the power spectrum, so that

$$
\mathcal{P}(p) = \mathcal{P}_{\mathrm{sm}}(p) \prod_k \mathcal{P}_{\mathrm{IG}}(p_k). \quad (23)
$$

As an example, we choose a smoothness-enforcing prior of the shape

$$
\mathcal{P}_{\mathrm{sm}}(p) \propto \exp \left(-\frac{1}{2\sigma_p^2} \int \mathrm{d}(\log k) \left(\frac{\partial^2 \log P_k}{\partial (\log k)^2} \right)^2 \right). \quad (24)
$$

This prior is maximized by any power-law power spectrum and punishes deviations from such a shape. The strength of the punishment is determined by the parameter σ_p. Other useful shapes for a smoothness prior could contain the first logarithmic derivative, punishing steep spectra, or simple derivatives with respect to k, punishing abrupt changes in the power spectrum. To illustrate the meaning of such smoothness priors and point out possible caveats, we discuss a few specific cases in App. A.

The spectral smoothness prior, Eq. (24), can be written as a Gaussian in $p = \log P$,

$$
\mathcal{P}_{\mathrm{sm}}(p) \propto \exp \left(-\frac{1}{2} p^\dagger T p \right), \quad (25)
$$

where the linear operator T includes both the second derivative and the scaling constant σ_p^2. The detailed form of the operator T that we use in our calculations can be found in App. B.

Introducing this prior corresponds to adding the term $\frac{1}{2} p^\dagger T p$ to the Hamiltonian in Eq. (20). Taking its derivative with respect to one p_k and equating the result with zero, we now get

$$
\mathrm{e}^{p_k} = \frac{q_k + \frac{1}{2} \sum_{\{\vec{k}'|k'=k\}} \left(\left| m_{\vec{k}'} \right|^2 + D_{\vec{k}'\vec{k}'} \right)}{\alpha_k - 1 + \frac{\rho_k}{2} + (Tp)_k}. \quad (26)
$$

The only point in which this equation differs from Eq. (12) is the extra term Tp in the denominator. As before, we can calculate the inverse Hessian of the Hamiltonian as an approximate uncertainty matrix. The Hessian is

$$
\begin{aligned}
\left. \frac{\partial^2 H(d, p)}{\partial p_k \partial p_{k'}} \right|_{p=p^{(\mathrm{MAP})}} &= \left(\alpha_k - 1 + \frac{\rho_k}{2} + (Tp)_k \right) \delta_{kk'} + T_{kk'} \\
&\quad - \frac{1}{2} \mathrm{e}^{-p_k - p_{k'}} \sum_{\substack{\{\vec{q}|q=k\} \\ \{\vec{q}'|q'=k'\}}} \left(2 \, \Re \big(m_{\vec{q}} m_{\vec{q}'}^* D_{\vec{q}\vec{q}'} \big) + |D_{\vec{q}\vec{q}'}|^2 \right) \bigg|_{p=p^{(\mathrm{MAP})}}
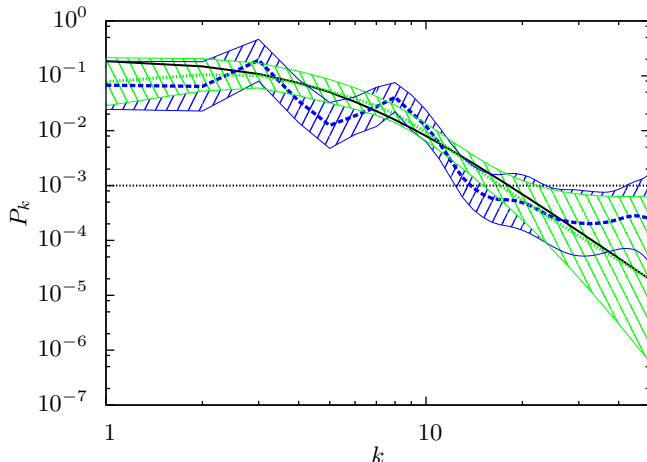\end{aligned}
$$

$$
(27)
$$

FIG. 3. Power spectra of the one-dimensional test-case with smoothness prior. The black solid line shows the theoretical power spectrum, the blue dashed line the power spectrum reconstructed with a smoothness prior with $\sigma_p^2 = 1000$, and the green dotted line the reconstruction with a stiffer smoothness prior with $\sigma_p^2 = 10$. The hatched regions around the dashed and dotted lines are the corresponding one-sigma uncertainty regions estimated from the inverse Hessian of the Hamiltonian. The horizontal dotted line indicates the noise level.

C.   Test-cases

Using the same one-dimensional test-case as shown in Figs. 1 and 2, we perform a reconstruction using the spectral smoothness prior given in Eq. (24). The resulting power spectra using a moderate strength for the prior with $\sigma_p^2 = 1000$ and a strict smoothness prior with $\sigma_p^2 = 10$ are shown in Fig. 3. Clearly, the new reconstructions are a much better approximation to the shape of the theoretical power spectrum than the one shown in Fig. 1. Also, as expected, the power spectrum obtained when using the strict smoothness prior turns out smoother and more closely resembles a power law.

Also shown in Fig. 3 is an uncertainty estimate for the reconstructed power spectra, obtained from the Hessian of the Hamiltonian. Taking a one-sigma uncertainty estimate for the logarithmic power spectrum components,

$$\Delta_p = \left[\mathrm{diag}\left(\left(\frac{\delta^2 H(d,p)}{\delta p \, \delta p^\dagger} \bigg|_{p=p^{(\mathrm{MAP})}} \right)^{-1} \right) \right]^{1/2}, \quad (28)$$

we plot the uncertainty interval for the power spectra as $P_k = \exp\left(p_k^{(\mathrm{MAP})} \pm \Delta_p \right)$. The uncertainty interval of the power spectrum estimates calculated in this way can, however, be misleading. It should be noted that the uncertainty of the power on the different k-modes is correlated. To illustrate this, we plot the complete uncertainty matrix, i.e. the inverse Hessian of the Hamiltonian, for the case with $\sigma_p^2 = 1000$ in Fig. 4. It can be seen from this figure that the correlation of the uncertainty on different k-modes is especially large for small scales. Furthermore,
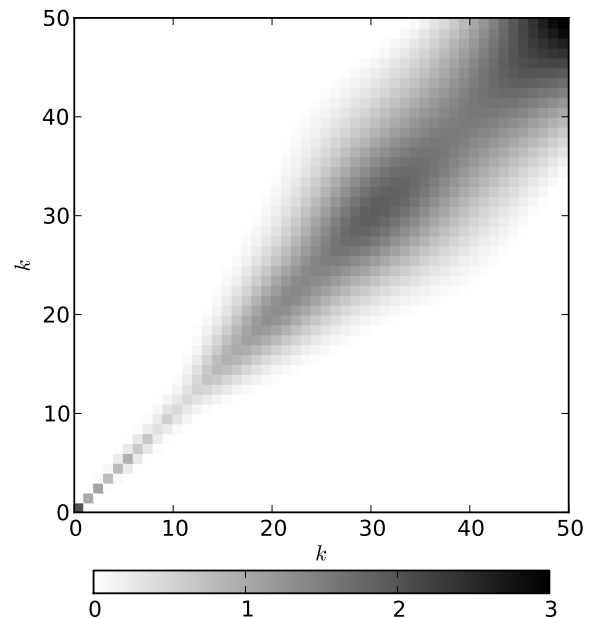

FIG. 4. Inverse Hessian of the Hamiltonian with a smoothness prior with $\sigma_p^2 = 1000$. Plotted is the full matrix given by the inverse of Eq. (27), evaluated at $p = p^{(\mathrm{MAP})}$. The diagonal of this matrix can be translated into the uncertainty interval plotted in Fig. 3, however, that does not take into account the correlations across different k-modes that are visible in this plot.

the inverse Hessian is only a rough approximation to the uncertainties in the power spectrum estimation, which typically underestimates these.

In all following examples we will use a spectral smoothness prior with an intermediate stiffness, given by $\sigma_p^2 = 100$.

To study the improvement that the spectral smoothness prior brings for the reconstruction of the signal field, we draw 100 different signal realizations from the same power spectrum and add 100 different noise realizations. For each of these data sets, we perform the full reconstruction and then calculate the power of the difference between the reconstructed field m and the true signal field s. In Fig. 5 we plot this power, averaged over the 100 realizations, for different reconstruction schemes. Under the assumption that the correct power spectrum is known a priori, the residual power is essentially given by the noise level on scales for which the signal is dominating and by the signal power on scales for which the noise is dominating. As can be seen from the plot, reconstructing the power spectrum without any spectral smoothing leads to a significantly increased residual power on the noise-dominated scales, while the quality of the reconstructions including the smoothness prior is close to the one of the Wiener filter reconstructions, showing only a slight excess in residual power on the smallest scales. We also plot the average residual power for reconstructions in
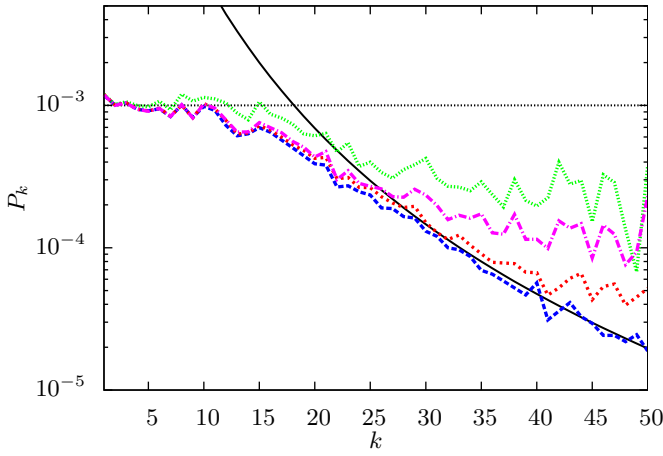
FIG. 5. Residual power between true signals and their reconstructions averaged over 100 different realizations. The black solid line shows the theoretical power spectrum of the signals, given by Eq. (14) with $P_0 = 0.2$, $k_0 = 5$, and $\gamma = 4$. The dotted horizontal line indicates the noise level, corresponding to $\sigma_n^2 = 0.1$. The remaining lines show the residual power for different reconstruction schemes. From top to bottom, these are a reconstruction without any spectral smoothing (green dotted line), with an ad-hoc convolution of the power spectrum (magenta dash-dotted line), with the smoothness prior given by Eq. (24) with $\sigma_p^2 = 100$ (red short-dashed line), and using the correct power spectrum (blue long-dashed line).
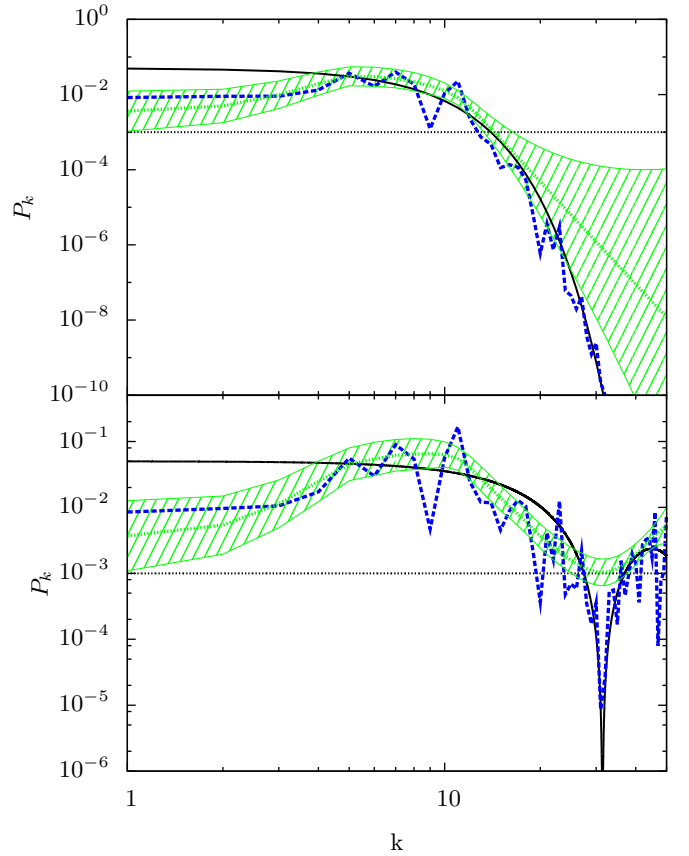

FIG. 6. Power spectra of signals with correlation functions given by Eq. (A4) (top panel) and Eq. (A6) (bottom panel). The black solid lines show the theoretical power spectra, the blue dashed lines the power in the random field realization drawn from the theoretical power spectra, and the green dotted lines the reconstructed power spectra using the spectral smoothness prior given in Eq. (24) with $\sigma_p^2 = 100$. The parameters substituted into Eq. (A4) are $\sigma = 0.2$ and $C_0 = 1/(4\sqrt{2\pi})$ and the ones used in Eq. (A6) are $L = 0.2$ and $C_0 = 0.25$. The noise variance in both cases is $\sigma_n^2 = 0.1$. The hatched area indicates the uncertainty of the reconstructed power spectrum estimated from the inverse Hessian and the horizontal dotted line the noise level.

which an ad-hoc smoothing step is applied to the power spectrum after each iteration of Eq. (12). This is implemented as a simple convolution with a Blackman window of width $\Delta_k = 9$. As can be seen in Fig. 5, this ad-hoc smoothing partly alleviates the problems of the power spectrum reconstruction but is clearly outperformed by the rigorous application of a smoothness prior.

Finally, we consider signal fields of the types discussed in App. A 2 and A 3, i.e. signals with Gaussian and triangular correlation functions given by Eqs. (A4) and (A6), respectively. As discussed in the appendix, these correlation functions lead to theoretical power spectra that are strongly suppressed by our spectral smoothness prior. Here, we investigate how serious this unwanted effect is in practice.

Shown in Figs. 6 and 7 are the results of the reconstruction for the power spectra and signal fields, respectively. We again use signals defined on an interval of length one, which we divide into 100 pixels, and choose $\sigma = 0.2$ and $C_0 = 1/(4\sqrt{2\pi})$ for the Gaussian correlation function, Eq. (A4), and $L = 0.2$ and $C_0 = 0.25$ for the triangular correlation function, Eq. (A6). In both cases, the noise variance is $\sigma_n^2 = 0.1$.

As expected, the features in the power spectra that are discouraged by the spectral smoothness prior are not well reconstructed. From Fig. 6 it is obvious that in the case with a Gaussian correlation function, the reconstructed and true power spectra deviate quite strongly on small scales. While the true power spectrum drops off rapidly, the reconstructed one stays comparatively flat. In the

case of a triangular correlation function, the same can be said about the finite k-value where the true power spectrum drops to zero. While the drop is indeed represented by the signal realization, as can be seen from the dashed line in the lower panel, the reconstructed power spectrum stays relatively level.

However, the features in the power spectra that the filter fails to reconstruct are below the noise level, whereas the rise of the power spectrum in the bottom panel of Fig. 6 on the smallest scales, which is above the noise level, is indeed represented in the reconstructed power spectrum as well. Thus, the signal reconstruction does not suffer significantly, as can be seen in Fig. 7. An accurate reconstruction of the power spectrum is important mainly at the scales on which the signal-response and noise are of comparable magnitude. If the signal power is
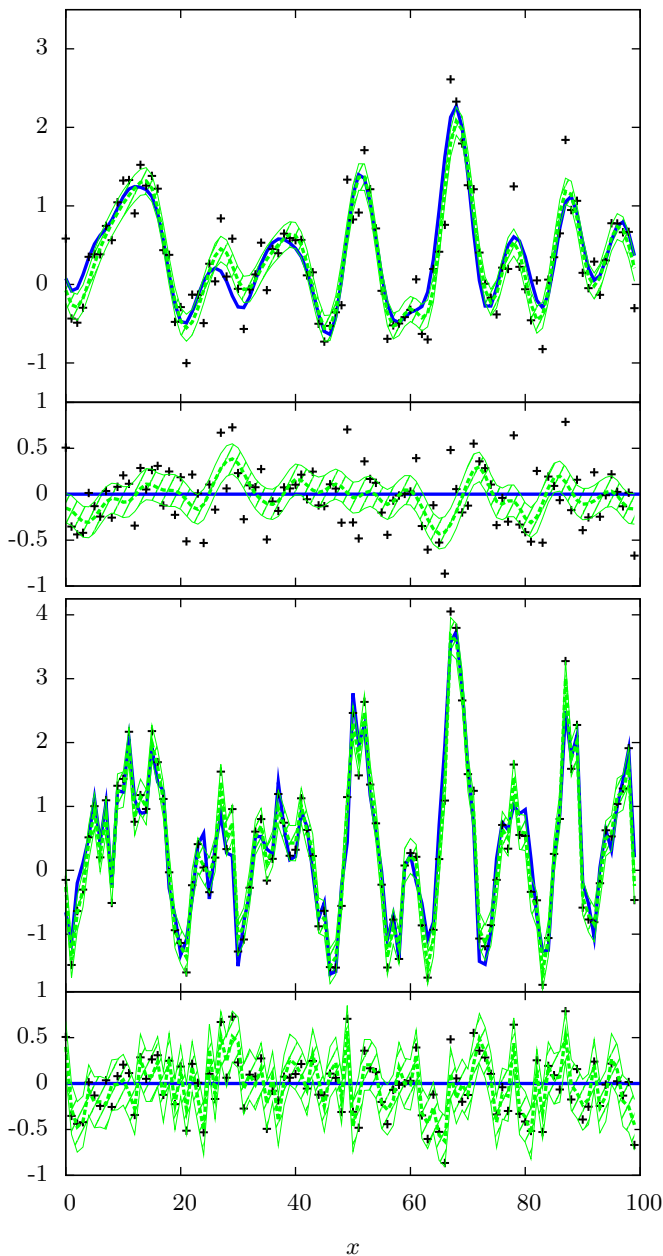
FIG. 7. Signal fields corresponding to the power spectra shown in Fig. 6. The top panel shows a signal with a Gaussian two-point correlation function, Eq. (A4), the bottom panel shows one with a triangular correlation function, Eq. (A6). Shown in each plot are the signal realization s (blue solid line), the data d (crosses), the signal reconstruction m (green dashed line), and an estimate of the local one-sigma uncertainty of the reconstruction, given by $m \pm \mathrm{diag}(D)^{1/2}$ (hatched area). At the bottom of each panel, the same is plotted with the true signal subtracted.

dominant, the reconstruction will follow the data closely, while it will smooth the data heavily in the opposite case of dominating noise, irrespective of the exact shape of the power spectrum. Thus, the reconstruction algorithm

will in general perform well even in cases in which certain features in the power spectra are not allowed by the spectral smoothness prior. However, if the objective is an accurate reconstruction of the power spectrum and any such features are suspected to be present, the spectral smoothness prior needs to be adapted to this situation.

III. RECONSTRUCTING LOG-NORMAL FIELDS

Now we turn to the problem of reconstructing a log-normal field. We define our signal field s to be the logarithm of this log-normal field, so that the prior probability distribution for s is again a Gaussian, described by a mean and a covariance S. For simplicity, we assume the prior mean to be zero. The data model then becomes

$$d = R\mathrm{e}^s + n, \tag{29}$$

where we have again included additive Gaussian noise n and the application of the exponential function to the signal field is to be interpreted pointwise. Due to the non-linearity of the exponential function, the posterior of s,

$$\mathcal{P}(s|d, S) \propto \mathcal{G}(d - R\mathrm{e}^s, N)\, \mathcal{G}(s, S), \tag{30}$$

is highly non-Gaussian, even when the signal covariance S is known. Adding a prior for S and marginalizing over it only makes the problem more complex,

$$\mathcal{P}(s|d) \propto \mathcal{G}(d - R\mathrm{e}^s, N) \int \mathcal{D}S\, \mathcal{G}(s, S)\, \mathcal{P}(S). \tag{31}$$

The path we pursue here is to treat this probability distribution approximatively. In order to do so, we employ the formalism of minimum Gibbs free energy presented in [8]. The basic idea is to approximate the posterior $\mathcal{P}(s|d)$ with a Gaussian, described by a mean m and a covariance D. An approximate Gibbs free energy can be calculated as a function of these quantities and it was shown in [8] that the optimal Gaussian approximation – in the sense of minimum Kullback-Leibler divergence – can be obtained by minimizing the approximate Gibbs free energy with respect to m and D.

The approximate Gibbs free energy is defined as

$$\tilde{G}(m, D) = \tilde{U}(m, D) - TS_{\mathrm{B}}(D). \tag{32}$$

The last term in this definition is the Boltzmann entropy

$$S_{\mathrm{B}}(D) = \frac{1}{2}\mathrm{tr}(1 + \log{(2\pi D)}), \tag{33}$$

which depends only on the covariance D. The other term in the approximate Gibbs free energy is the approximate internal energy

$$\tilde{U}(m, D) = \langle H(s, d) \rangle_{\mathcal{G}(s - m, D)}. \tag{34}$$

Here, $\langle \cdot \rangle_{\mathcal{G}(s-m,D)}$ denotes an expectation value calculated with respect to the Gaussian posterior approximation and $H(s,d) = -\log \mathcal{P}(s,d)$ is the Hamiltonian of the problem. Finally, the temperature T can be regarded as a tuning parameter that regulates the importance that is given to the region around the maximum of the full posterior and regions removed from the maximum. The limiting value $T = 0$ leads to the maximum a posteriori solution for the signal field. In the following, we choose the default value of $T = 1$ and refer the reader to [8, 9, 11] for an in-depth discussion of the temperature parameter.

Assuming independent inverse-gamma distributions as priors for the power spectrum components P_k, as we had already done in Eq. (17), the calculation of the Hamiltonian yields

$$
\begin{aligned}
H(s,d) &= -\log\left(\mathcal{P}(d|s)\,\mathcal{P}(s|P)\,\mathcal{P}(P)\right) \\
&= -\log\left(\mathcal{G}(d - Re^s, N)\,\mathcal{G}(s, S)\prod_k \mathcal{P}_{\text{IG}}(P_k)\right) \\
&= -j^\dagger e^s + \frac{1}{2}\left(e^s\right)^\dagger M e^s \\
&\quad + \sum_k \left(\alpha_k - 1 + \frac{\rho_k}{2}\right)\log\left(q_k + \frac{1}{2}\sum_{\{\vec{k}'|k'=k\}}\left|s_{\vec{k}'}\right|^2\right) \\
&\quad + \text{const.}
\end{aligned}
\tag{35}
$$

Here, we have again collected s-independent terms in an additive constant and introduced the abbreviation

$$
M = R^\dagger N^{-1} R. \tag{36}
$$

In order to calculate the Gaussian expectation value of this Hamiltonian analytically, we expand the logarithm appearing in this expression in a power series around its expectation value

$$
\begin{aligned}
\tilde{q}_k &= \left\langle q_k + \frac{1}{2}\sum_{\{\vec{k}'|k'=k\}}\left|s_{\vec{k}'}\right|^2\right\rangle_{\mathcal{G}(s-m,D)} \\
&= q_k + \frac{1}{2}\sum_{\{\vec{k}'|k'=k\}}\left(\left|m_{\vec{k}'}\right|^2 + D_{\vec{k}'\vec{k}'}\right)
\end{aligned}
\tag{37}
$$

as

$$
\begin{aligned}
&\log\left(q_k + \frac{1}{2}\sum_{\{\vec{k}'|k'=k\}}\left|s_{\vec{k}'}\right|^2\right) \\
&\approx \log\left(\tilde{q}_k\right) \\
&\quad - \sum_{i=1}^{i_{\max}}\frac{(-1)^i}{i\tilde{q}_k^i}\left(q_k + \frac{1}{2}\left(\sum_{\{\vec{k}'|k'=k\}}\left|s_{\vec{k}'}\right|^2\right) - \tilde{q}_k\right)^i.
\end{aligned}
\tag{38}
$$

We truncate this expansion after the first order, i.e. $i_{\max} = 1$. Note that our choice of $\tilde{q}_k$ ensures that the first-order term itself vanishes.

With this simplification, we can calculate the approximate Gibbs free energy to be

$$
\begin{aligned}
\tilde{G}(m, D) &= \int_{\mathcal{M}}\mathrm{d}x\, j_x e^{m_x + \frac{1}{2}D_{xx}} \\
&\quad + \int_{\mathcal{M}}\mathrm{d}x\int_{\mathcal{M}}\mathrm{d}y\,\frac{1}{2}M_{xy}e^{m_x + m_y + \frac{1}{2}D_{xx} + \frac{1}{2}D_{yy} + D_{xy}} \\
&\quad + \sum_k\left(\alpha_k - 1 + \frac{\rho_k}{2}\right)\log\left(\tilde{q}_k\right) \\
&\quad - \frac{1}{2}\mathrm{tr}(1 + \log\left(2\pi D\right)).
\end{aligned}
\tag{39}
$$

To avoid confusion, we write out explicitly all integrals appearing here and in the following. Taking the functional derivatives with respect to m and D and equating them with zero yields two filter equations that determine m and D. In these equations, the right hand side expression of Eq. (12) appears. If we reidentify this expression with the spectral components P_k and write $S_{\vec{k}\vec{k}'} = \delta_{\vec{k}\vec{k}'}P_k$, the two filter equations become

$$
\begin{aligned}
m_x &= \int_{\mathcal{M}}\mathrm{d}y\, S_{xy}\left[j_y e^{m_y + \frac{1}{2}D_{yy}}\right. \\
&\quad \left. - \int_{\mathcal{M}}\mathrm{d}z\, M_{yz}e^{m_y + m_z + \frac{1}{2}D_{yy} + \frac{1}{2}D_{zz} + D_{yz}}\right]
\end{aligned}
\tag{40}
$$

and

$$
\begin{aligned}
\left(D^{-1}\right)_{xy} &= -j_x e^{m_x + \frac{1}{2}D_{xx}}\delta_{xy} \\
&\quad + \int_{\mathcal{M}}\mathrm{d}z\, M_{xz}e^{m_x + m_z + \frac{1}{2}D_{xx} + \frac{1}{2}D_{zz} + D_{xz}}\delta_{xy} \\
&\quad + M_{xy}e^{m_x + m_y + \frac{1}{2}D_{xx} + \frac{1}{2}D_{yy} + D_{xy}} \\
&\quad + \left(S^{-1}\right)_{xy}.
\end{aligned}
\tag{41}
$$

Together with Eq. (12), the last two equations fully determine the Gaussian approximation to the posterior. Solving them self-consistently gives an estimate m for the signal field and an estimate D for the corresponding uncertainty matrix. The corresponding approximate posterior mean estimate of the exponentiated field values then is

$$
\langle e^{s_x}\rangle_{\mathcal{P}(s|d)} \approx \langle e^{s_x}\rangle_{\mathcal{G}(s-m,D)} = e^{m_x + \frac{1}{2}D_{xx}}. \tag{42}
$$

Of special interest for many applications is the case in which the matrix $M = R^\dagger N^{-1}R$ is diagonal, e.g. when the noise contributions to the individual data points are uncorrelated and the response is purely local. In this case the filter equations simplify somewhat to

$$
m_x = \int_{\mathcal{M}}\mathrm{d}y\, S_{xy}\left[j_y e^{m_y + \frac{1}{2}D_{yy}} - M_{yy}e^{2m_y + 2D_{yy}}\right] \tag{43}
$$

and

$$
\begin{aligned}
\left(D^{-1}\right)_{xy} &= \left[-j_x e^{m_x + \frac{1}{2}D_{xx}} + 2M_{xx}e^{2m_x + 2D_{xx}}\right]\delta_{xy} \\
&\quad + \left(S^{-1}\right)_{xy}.
\end{aligned}
\tag{44}
$$

A. Spectral smoothness in the log-normal case

In Sec. II A we had seen that in the Gaussian case the full reconstruction with unknown power spectrum can be regarded as a combination of the posterior mean reconstruction under the assumption of a known power spectrum and the estimation of this power spectrum as the one that maximizes its posterior probability. In the case of a log-normal field, the posterior mean cannot be calculated analytically, even if the power spectrum is assumed to be known. However, it is interesting to note that employing the formalism of minimum Gibbs free energy to the log-normal reconstruction problem with known power spectrum, one arrives exactly at the formulas derived in the previous section, Eqs. (40) and (41).

Thus, the full reconstruction, consisting of Eqs. (40), (41), and (12) can again be regarded as a combination of the calculation of the posterior mean for the signal under the assumption of a power spectrum and the estimation of the power spectrum according to Eq. (12). From this viewpoint, the inclusion of a smoothness prior for the power spectrum is trivial. We simply replace the power spectrum estimation step according to Eq. (12) with the one derived in Sec. II B, i.e. with Eq. (26), just as we had done in the Gaussian case.

B. Test cases

In this section, we present some test cases for the theory developed so far on the reconstruction of log-normal fields. We study one-dimensional tests with differing degrees of non-linearity and differing noise-levels, as well as two two-dimensional test-cases. For simplicity, we assume an ideal local response, $R = \mathbb{1}$, except in the last example, where we study the effects of an observational mask. The noise is assumed to be uncorrelated and homogeneous, $N = \sigma_n^2 \mathbb{1}$. For the smoothness prior, we use the one discussed in Sec. II B and choose $\sigma_p^2 = 100$.

First, we discuss again the case of a signal on the one-sphere $\mathcal{S}^1$, i.e. the interval $[0, 1)$ with periodic boundary conditions, here discretized into 100 pixels. Shown in Figs. 8-10 are three different cases. In each of these cases, the signal realization is the same as in Sec. II, only with a differing normalization of the power spectrum. In Fig. 8 we show a mildly non-linear case with $P_0 = 0.1$ and in Figs. 9 and 10 a highly non-linear case with $P_0 = 0.3$. The noise variance σ_n^2 is 0.1 in Fig. 8, 1 in Fig. 9, and 25 in Fig. 10. In the lower panels of these three figures, we show the signals and their reconstructions and in the upper panels the exponentiated fields, e^s, as well as our posterior mean estimate for these, $\langle e^s \rangle_{\mathcal{P}(s|d)} \approx e^{m + \frac{1}{2} \mathrm{diag}(D)}$.

In the mildly non-linear case the reconstruction is a reasonably good approximation to the true signal both in regions of high signal values and regions of low signal values. However, it is apparent that the quality of the reconstruction is slightly higher in the former regions. This is due to the homogeneity of the noise statistics that
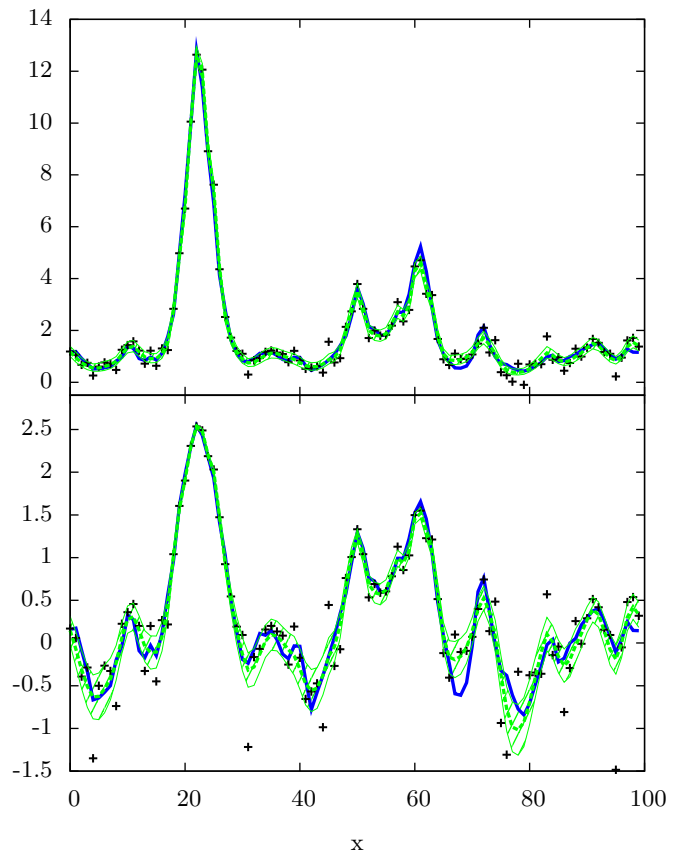

FIG. 8. One-dimensional log-normal reconstruction in a mildly non-linear low-noise case. The power spectrum from which the signal realization was drawn is given by Eq. (14) with $P_0 = 0.1$, $k_0 = 5$, and $\gamma = 4$. The noise variance is $\sigma_n^2 = 0.1$. The top panel shows the exponentiated signal field e^s (blue solid line), the data d (crosses), the reconstruction $e^{m + \frac{1}{2} \mathrm{diag}(D)}$ (green dashed line), and the uncertainty interval of the reconstruction, given by $e^{m + \frac{1}{2} \mathrm{diag}(D) \pm (\mathrm{diag}(D))^{1/2}}$ (hatched region). The lower panel shows the signal field s (blue solid line), the logarithm of the data $\log(d)$ (crosses), the reconstruction m (green dashed line), and the uncertainty interval for m, given by $m \pm (\mathrm{diag}(D))^{1/2}$. In the lower panel, only data points for which $\log d$ is greater than -1.5 are shown.

we have assumed. Since our signal response, given by Re^s, depends non-linearly on the signal, the noise impact is lower in regions of higher signal values and hence the signal inference is less demanding in these regions. From Fig. 8 it is apparent that this effect is well represented by the point-wise uncertainty of the reconstruction, given by $\mathrm{diag}(D)$. As can be seen from Fig. 9, the effect becomes more pronounced for higher degrees of non-linearity, i.e. larger signal variances on a logarithmic scale.

The highly nonlinear case with high noise level, depicted in Fig. 10, exhibits a point-wise uncertainty estimate for the reconstruction that can clearly not be interpreted as a 68% confidence interval. This is due to a known problem of the filter discussed in Sec. II and by
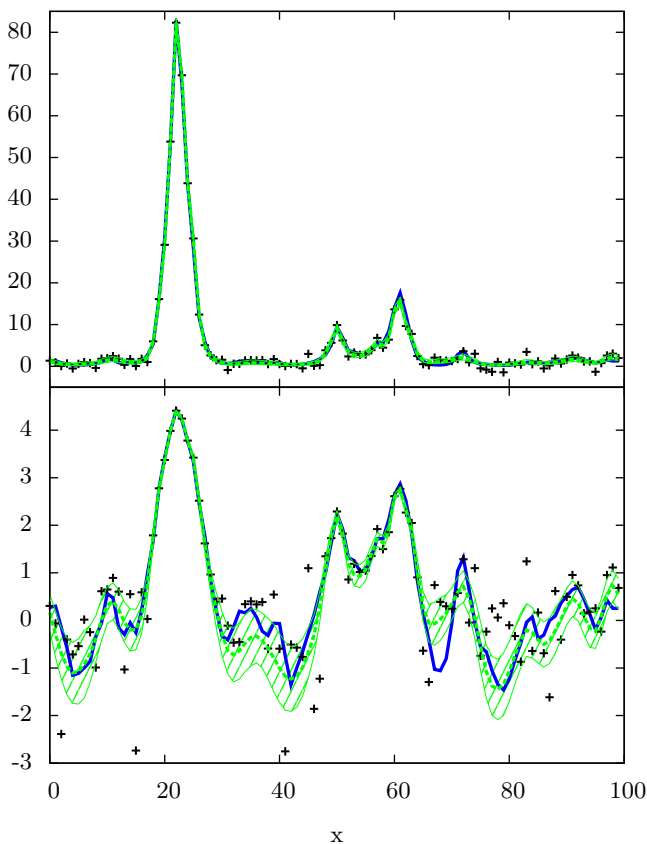
FIG. 9. One-dimensional log-normal reconstruction in a highly non-linear low-noise case. The plotted quantities are the same as in Fig. 8. The parameters describing the power spectrum are $P_0 = 0.3$, $k_0 = 5$, and $\gamma = 4$ and the noise variance is $\sigma_n^2 = 1$. In the lower panel, only data points for which $\log d$ is greater than -3 are shown.
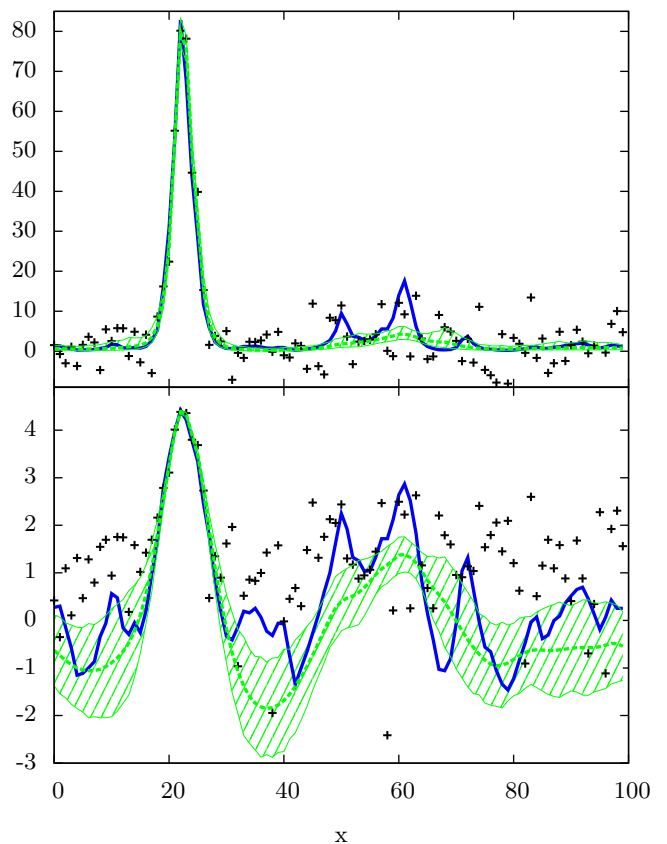
FIG. 10. One-dimensional log-normal reconstruction in a highly non-linear high-noise case. The plotted quantities are the same as in Fig. 8. The parameters describing the power spectrum are $P_0 = 0.3$, $k_0 = 5$, and $\gamma = 4$ and the noise variance is $\sigma_n^2 = 25$. In the lower panel, only data points for which $\log d$ is greater than -3 are shown.

extension also of the filter derived in this section. As discussed in [6], the filter exhibits a perception threshold. This means that if the signal-response-to-noise ratio is lower than a certain threshold on a given scale, then the filter will not reconstruct any power on this scale. Our usage of the spectral smoothness prior partly alleviates this porblem in that it prevents the power of individual scales to drop to zero. However, the reconstructed power on the noise-dominated scales will in general still be too low. This directly affects the estimate for the reconstruction's uncertainty, given by $D = \left(S^{-1} + M\right)^{-1}$, which tends toward zero in the limit of zero power on all scales.

Furthermore, the lack of power on all but the few signal-dominated scales can lead to ringing effects, i.e. prominent signal-dominated features are well reconstructed and extrapolated periodically into the noise-dominated regions. The deep trough in the signal reconstruction that can be seen in the lower panel of Fig. 10 around pixel number 35 is most likely due to this effect. The high degree of non-linearity acts to reinforce this effect. As can be seen in the top panel of Fig. 10, the ef-

fect of the trough onto the exponentiated reconstruction is almost intangible.

It is certainly mandatory to keep the potential problems of ringing and an underestimated uncertainty in mind when reconstructing a field which is swamped by noise in the better part of its domain. However, as can be seen from Fig. 10, the main features of the signal field will still be reconstructed reliably even in such an unfavorable case.

In these test cases, we have chosen the one-dimensional interval as domain of the log-normal signal field mainly for illustrative purposes. The algorithm is, however, versatile in that it can operate on virtually any space[1]. To illustrate this point, we close this section with two two-dimensional examples.

In Fig. 11, we show an example for a signal defined on the two-torus $\mathcal{T}^2 = \mathcal{S}^1 \times \mathcal{S}^1$, i.e. a segment of $\mathbb{R}^2$

————

[1] We are using the NIFTY package [23] in our implementation. This makes changing the domain of the signal field trivial, requiring only minuscule changes to the code.
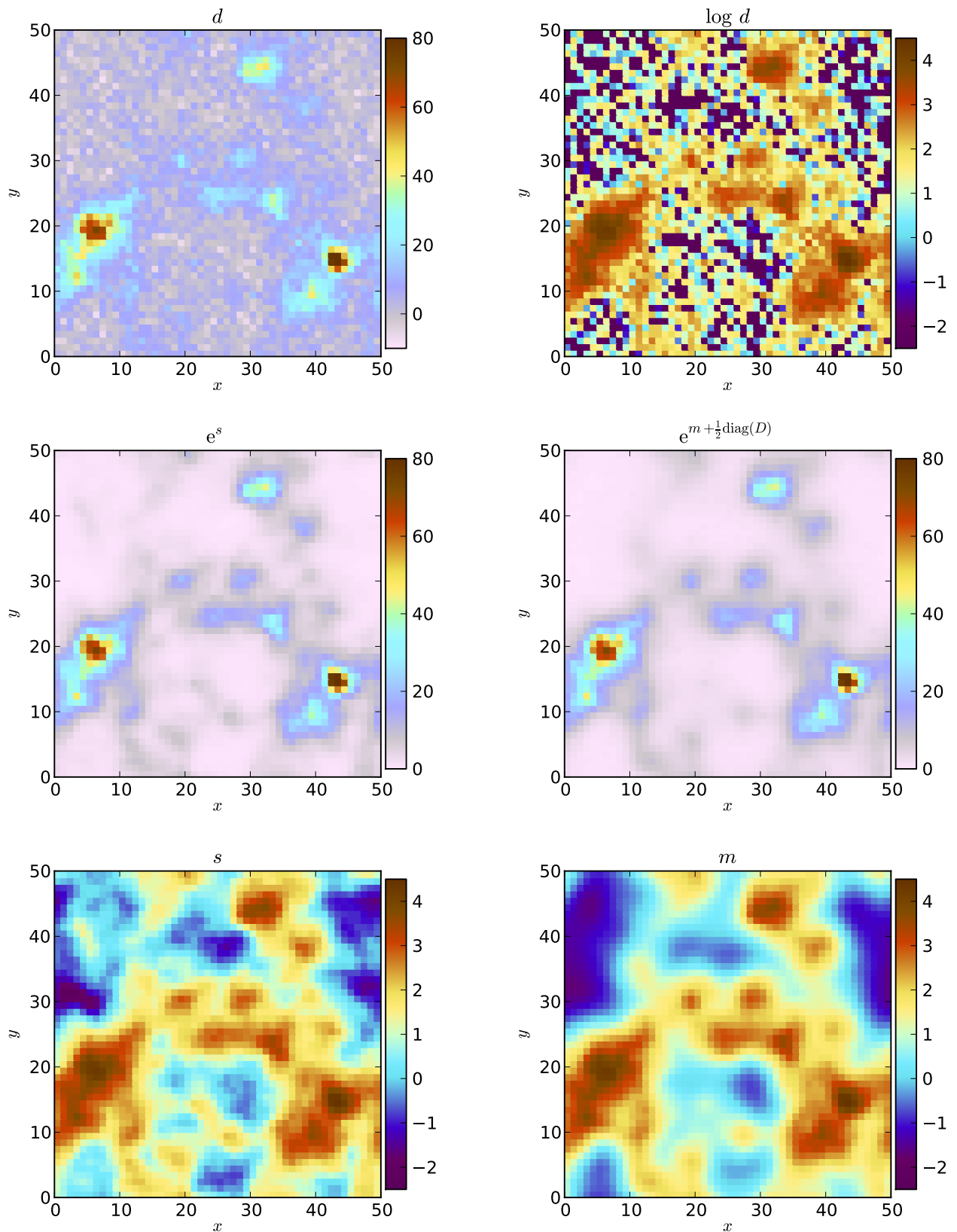
FIG. 11. Log-normal reconstruction example in a toroidal setting. The signal's power spectrum is given by Eq. (14) with $P_0 = 0.3$, $k_0 = 2$, and $\gamma = 5$. The noise variance is $\sigma_n^2 = 10$. The top row shows the data set d (left) and its logarithm (right). In the logarithmic version, pixels with negative data values are plotted as dark blue. The middle row shows the exponentiated signal field e^s (left) and its reconstruction, given by $\mathrm{e}^{m+\frac{1}{2}\mathrm{diag}(D)}$, (right). The corresponding non-exponentiated quantities s (left) and m (right) are shown in the bottom row.
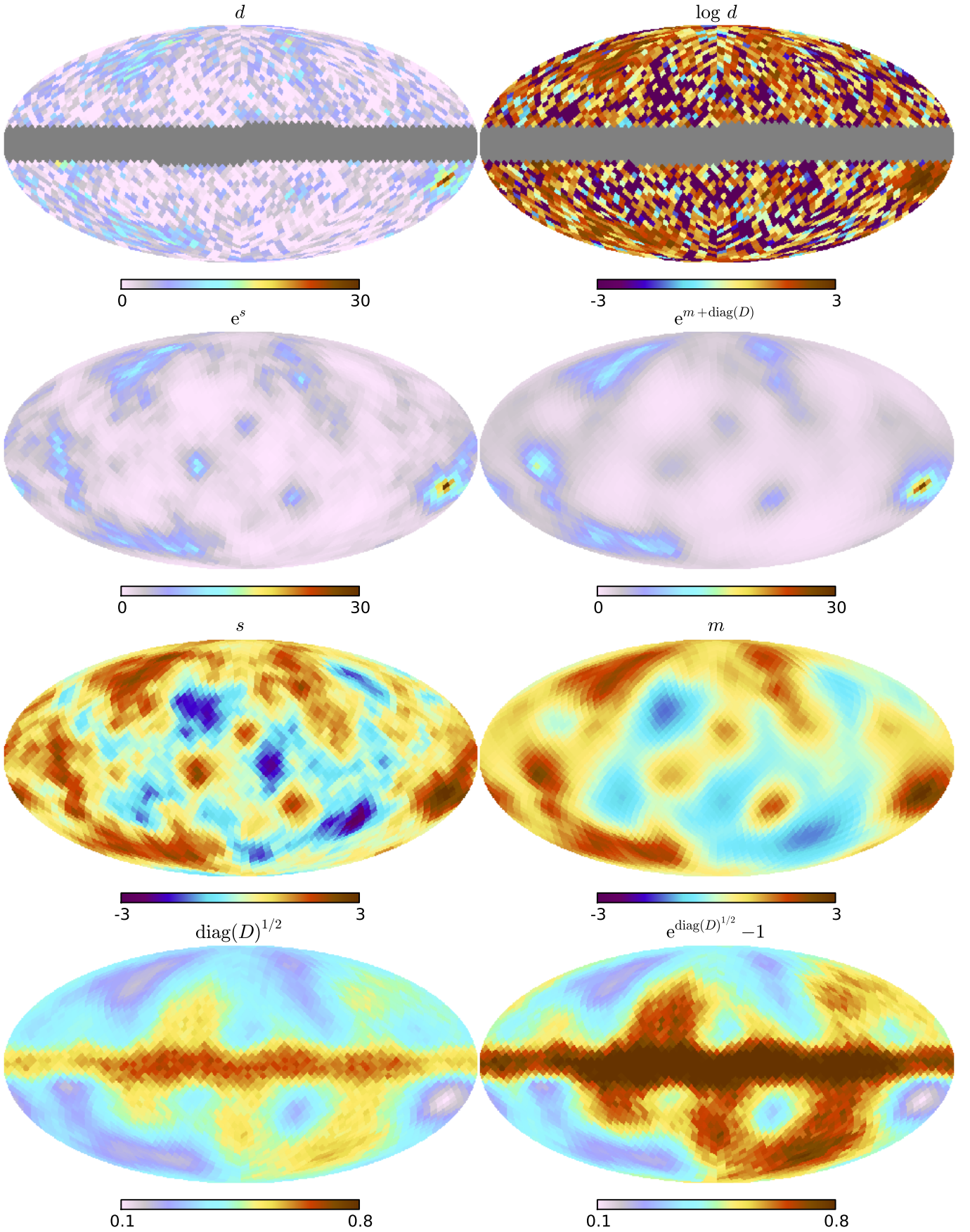
FIG. 12. Log-normal reconstruction example on a two-sphere. The quantities plotted in the top three rows are the same as in Fig. 11. The bottom row shows the uncertainty of the signal reconstruction given by $\mathrm{diag}(D)^{1/2}$ (left side) and the fractional uncertainty of the reconstruction of the exponentiated field given by $\mathrm{e}^{\mathrm{diag}(D)^{1/2}} - 1$ (right side). The power spectrum parameters are $P_0 = 0.3$, $k_0 = 5$, and $\gamma = 4$ and the noise variance is $\sigma_n^2 = 10$. In the top row, pixels without measurements are plotted gray.

with periodic boundary conditions. The periodicity here and in the earlier one-dimensional examples is a consequence of our use of Fast Fourier Transform routines. It can be approximately overcome by artificially extending the interval or the section of $\mathbb{R}^2$ for more than the typical correlation length of the signal while leaving the data unchanged when doing the reconstruction. Thus, the influence of the reconstruction on one side of the interval or rectangle will have negligible influence on the reconstruction on the other side.

We discretize the two-torus and its corresponding Fourier space into 50×50 pixels. The rectangular orientation of the Fourier pixels leads to many different scales k being represented, each, however, only by a few pixels, corresponding to a few different Fourier vectors $\vec{k}$. To avoid dealing with all these scales individually, we bin the scales logarithmically into 17 bins. We replace all summations over Fourier components with a certain scale appearing in the filter formulas with summations over all Fourier components whose scale falls within one bin. Thus, we reconstruct the power for each bin instead of for each scale that is represented in the rectangular Fourier grid. The signal is again drawn from the power spectrum given by Eq. (14) with $P_0 = 0.3$, $k_0 = 2$, and $\gamma = 5$. The noise level is chosen as $\sigma_n^2 = 10$. Signal, data, and reconstruction are shown in Fig. 11.

The toroidal example exhibits essentially the same features that we had already seen in the one-dimensional case. While the reconstruction is generally in good agreement with the true underlying signal, it is less accurate in the regions of small signal values where the signal-response-to-noise ratio is far worse due to the quite high degree of non-linearity.

In Fig. 12, we show an example for a reconstruction on the two-sphere $\mathcal{S}^2$. This example has special relevance for astronomical applications since one can interpret the sky as a two-sphere and therefore any astrophysical signals without distance information will be defined on this manifold. We generate a mock signal from the power spectrum given in Eq. (14) with $P_0 = 0.3$, $k_0 = 5$, and $\gamma = 4$. The noise level is $\sigma_n^2 = 10$. We discretize the sphere using the HEALPIX[2] package [10] with a resolution parameter of $N_{\text{side}} = 16$ leading to $3\,072$ pixels in total. The power spectrum components P_k are in this case the components of an angular power spectrum, often denoted as C_ℓ.

In this last case study, we replace the trivial response $R = \mathbb{1}$ with a projection onto part of the sphere, effectively masking 600 pixels around the equator for which we assume that no measurements have been taken. This resembles a typical situation in extragalactic astronomy, where observations through the Galactic plane are not possible due to the obscuration by the Milky Way.

─────────

[2] The HEALPIX package is available from `http://healpix.jpl.nasa.gov/`.

Fig. 12 shows the signal field, data, and reconstruction both in the exponentiated and in the linear version. In the panels showing the reconstructed signal field, it can be nicely seen how the algorithm is able to extrapolate into the gap region from the data on the boundary. This is possible due to the knowledge of the correlation structure that was inferred from the same data set. Also shown in Fig. 12 are the pixel-wise uncertainty estimate of the signal field's reconstruction, given by $\text{diag}(D)^{1/2}$, and the fractional uncertainty of the reconstructed exponentiated signal, given by $e^{\text{diag}(D)^{1/2}} - 1$, which is approximated well by $\text{diag}(D)^{1/2}$ in most regions. It can be seen that the uncertainty tends to be higher in the regions of low signal values, as was the case in the one-dimensional examples, and in the region around the equator which lacks observations. This is to be expected, since the only constraint on the signal in this region comes from extrapolations from neighboring regions using the signal's correlation structure inferred from the data.

IV. SUMMARY AND CONCLUSIONS

We have developed an algorithm to infer log-normal random fields from noisy measurement data. The log-normal model was chosen due to its wide range of applications in astrophysics and other fields. The reconstruction method uses the correlation structure of the log-normal field to differentiate between features in the data that are due to noise and such that are due to variations in the true underlying field. This correlation structure, determined by the field's power spectrum, is, however, in general not known a priori. We have therefore extended the theory for simultaneous reconstructions of a field and its power spectrum that was developed and applied successfully for Gaussian random fields in the past [6, 8, 19–21, 24] to log-normal fields.

An additional feature of our reconstruction method is the use of a smoothness prior for the power spectrum. We have suggested to employ a prior based on the second double-logarithmic derivative of the power spectrum and shown that it is well suited to handle a large variety of cases. Having investigated possible pitfalls associated with the usage of such a prior from theoretical as well as practical viewpoints, we should stress that the derivation of the filter formulas laid out here does not depend on the specific form of the spectral smoothness prior. In cases in which the prior we employed here cannot be expected to yield satisfactory results it should simply be replaced by a different one.

The algorithm we have derived depends in no way on the space on which the signal field is to be reconstructed. We have demonstrated this by showing examples of reconstructions of mock signals on a one-dimensional interval, a flat two-dimensional space, and a spherical space. We have discussed the performance of the algorithm in these scenarios and pointed out possible caveats when dealing with very low signal-response-to-noise ratios.

In these application examples, we have assumed that the observational data simply represent the underlying log-normal field, subject to additive noise. Furthermore, we have illustrated the ability of the algorithm to extrapolate from the given data in a test case in which these data were assumed to be incomplete, thus demonstrating the power of the usage of the correlation information contained in the data. The derivation of the filter formulas, however, is even more general. It allows for an arbitrary linear relationship between log-normal field and data, described by a response matrix R. The resulting formulas include the general response matrix which can e.g. represent an incomplete observation, a convolution, or a Fourier transformation. We have also allowed for a general noise covariance matrix, thus including cases of heteroscedastic or correlated noise.

This makes the algorithm widely applicable. Applications that we have in mind include for example the study of diffuse Galactic emission components at radio frequencies and reconstructions of emissivity fields across galaxy clusters from interferometric observations.

[1] Basu, S. & Dassios, A., A Cox process with log-normal intensity, Insurance: mathematics and economics **31** no. 2, (October 2002) 297–302, © 2002 Elsevier, http://eprints.lse.ac.uk/16375/

[2] Beneš, V., Bodlák, K., Møller, J., & Waagepetersen, R. P. 2002, Bayesian analysis of log Gaussian Cox processes for disease mapping, Tech. Rep. R-02-2001, Department of Mathematical Sciences, Aalborg University

[3] Brix, A., Space-time Multi Type Log Gaussian Cox Processes with a View to Modelling Weeds, Scandinavian Journal of Statistics **28** no. 3, (2001) 471–488, http://EconPapers.repec.org/RePEc:bla:scjsta:v:28:y:2001:i:3:p:471-488

[4] Brix, A. & Diggle, P. J., Spatiotemporal prediction for log-Gaussian Cox processes, Journal of the Royal Statistical Society Series B **63** no. 4, (2001) 823–841, http://EconPapers.repec.org/RePEc:bla:jorssb:v:63:y:2001:i:4:p:823-841

[5] Coles, P. & Jones, B., A lognormal model for the cosmological mass distribution, MNRAS **248** (Jan. 1991) 1–13, ADS

[6] Enßlin, T. A. & Frommert, M., Reconstruction of signals with unknown spectra in information field theory with parameter uncertainty, Phys. Rev. D **83** no. 10, (May 2011) 105014, arXiv:1002.2928 [astro-ph.IM]

[7] Enßlin, T. A., Frommert, M., & Kitaura, F. S., Information field theory for cosmological perturbation reconstruction and nonlinear signal analysis, Phys. Rev. D **80** no. 10, (Nov. 2009) 105005, arXiv:0806.3474

[8] Enßlin, T. A. & Weig, C., Inference with minimal Gibbs free energy in information field theory, Phys. Rev. E **82** no. 5, (Nov. 2010) 051112, arXiv:1004.2868 [astro-ph.IM]

[9] Enßlin, T. A. & Weig, C., Reply to "Comment on 'Inference with minimal Gibbs free energy in information field theory' ", Phys. Rev. E **85** no. 3, (Mar. 2012) 033102

[10] Górski, K. M., Hivon, E., Banday, A. J., et al., HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, ApJ **622** (Apr. 2005) 759–771, arXiv:astro-ph/0409513

[11] Iatsenko, D., Stefanovska, A., & McClintock, P. V. E., Comment on "Inference with minimal Gibbs free energy in information field theory", Phys. Rev. E **85** no. 3, (Mar. 2012) 033101

[12] Jasche, J., Kitaura, F. S., Li, C., & Enßlin, T. A., Bayesian non-linear large-scale structure inference of the Sloan Digital Sky Survey Data Release 7, MNRAS **409** (Nov. 2010) 355–370, arXiv:0911.2498 [astro-ph.CO]

[13] Jasche, J., Kitaura, F. S., Wandelt, B. D., & Enßlin, T. A., Bayesian power-spectrum inference for large-scale structure data, MNRAS **406** (Jul. 2010) 60–85, arXiv:0911.2493 [astro-ph.CO]

[14] Jasche, J. & Wandelt, B. D., Bayesian physical reconstruction of initial conditions from large scale structure surveys, ArXiv e-prints (Mar. 2012), arXiv:1203.3639 [astro-ph.CO]

[15] Kolmogorov, A., The Local Structure of Turbulence in Incompressible Viscous Fluid for Very Large Reynolds' Numbers, Akademiia Nauk SSSR Doklady **30** (1941) 301–305, ADS

[16] Lemm, J. 2003, Bayesian Field Theory, , Johns Hopkins University Press

[17] Møller, J., Syversveen, A. R., & Waagepetersen, R. P., Log Gaussian Cox Processes, Scandinavian Journal of Statistics **25** no. 3, (1998), http://dx.doi.org/10.2307/4616515

[18] Oh, S. P., Spergel, D. N., & Hinshaw, G., An Efficient Technique to Determine the Power Spectrum from Cosmic Microwave Background Sky Maps, ApJ **510** (Jan. 1999) 551–563, arXiv:astro-ph/9805339

[19] Oppermann, N., Junklewitz, H., Robbers, G., et al., An improved map of the Galactic Faraday sky, A&A **542** (Jun. 2012) A93, arXiv:1111.6186 [astro-ph.GA]

[20] Oppermann, N., Junklewitz, H., Robbers, G., & Enßlin, T. A., Probing magnetic helicity with synchrotron radiation and Faraday rotation, A&A **530** (Jun. 2011) A89, arXiv:1008.1246 [astro-ph.IM]

[21] Oppermann, N., Robbers, G., & Enßlin, T. A., Reconstructing signals from noisy data with unknown signal and noise covariance, Phys. Rev. E **84** no. 4, (Oct. 2011) 041118, arXiv:1107.2384 [astro-ph.IM]

[22] Pati, D., Reich, B. J., & Dunson, D. B., Bayesian geostatistical modelling with informative sampling locations, Biometrika **98** no. 1, (2011) 35–48, `http://EconPapers.repec.org/RePEc:oup:biomet:v:98:y:2011:i:1:p:35-48`

[23] Selig, M., e. a., in prep.

[24] Selig, M., Oppermann, N., & Enßlin, T. A., Improving stochastic estimates with inference methods: Calculating matrix diagonals, Phys. Rev. E **85** no. 2, (Feb. 2012) 021134, `arXiv:1108.0600 [astro-ph.IM]`

[25] Uhlenbeck, G. E. & Ornstein, L. S., On the Theory of the Brownian Motion, Phys. Rev. **36** (Sep 1930) 823–841, `http://link.aps.org/doi/10.1103/PhysRev.36.823`

[26] Wandelt, B. D., Larson, D. L., & Lakshminarayanan, A., Global, exact cosmic microwave background data analysis using Gibbs sampling, Phys. Rev. D **70** no. 8, (Oct. 2004) 083511, `arXiv:astro-ph/0310080`

Appendix A: Case studies for the spectral smoothness prior

In this appendix, we discuss a few possible correlation structures and power spectra and examine the suitability of a spectral smoothness prior of the form described by Eq. (24).

1. Power law spectra

If the signal's power spectrum is a broken power law of the shape given by Eq. (14), the first double-logarithmic derivative is

$$\frac{\partial p_k}{\partial \log k} = -\frac{\gamma}{1 + \left(\frac{k_0}{k}\right)^2}, \tag{A1}$$

which behaves like $-\gamma \left(\frac{k}{k_0}\right)^2$ for $k \ll k_0$ and tends toward $-\gamma$ for $k \to \infty$. The second double-logarithmic derivative is

$$\frac{\partial^2 p_k}{\partial \left(\log k\right)^2} = -\frac{2\gamma}{\left(\frac{k}{k_0} + \frac{k_0}{k}\right)^2}, \tag{A2}$$

which tends toward zero both for $k \to 0$ and $k \to \infty$. The second derivative takes on its extremum, given by $-\frac{\gamma}{2}$, at the knee frequency k_0 where the spectral index of the power spectrum changes from 0 to γ. Thus neither a prior punishing large values for the first logarithmic derivative, nor one punishing large values for the second logarithmic derivative prevents the true solution from being found, provided the values of $-\gamma$ and $-\frac{\gamma}{2}$, respectively, are well within the range of values that are allowed for the derivative by the prior. In fact, choosing $\sigma_p = \gamma/2$ would allow such a change in spectral index roughly once per e-folding of the k-value. Choosing $\sigma_p = \frac{\gamma}{2h}$ would turn such a kink into a less common h-sigma event.

The case of a broken power-law contains the special cases of a pure power-law, i.e. $k_0 \to 0$, such as arises for example for Brownian motion of a particle, and of an exponential two-point correlation function, i.e.

$$C(r) = S_{xy} = C_0 \mathrm{e}^{-\beta r}, \tag{A3}$$

where $r = |x - y|$ is the distance between two points or time-instances x and y. Such a correlation function arises for example from the Ornstein-Uhlenbeck process [25]. Fourier transforming this correlation function in a one-dimensional space yields the power spectrum which takes on the form given by Eq. (14) with $P_0 = \frac{2C_0}{\beta}$, $k_0 = \beta$, and $\gamma = 2$.

However, there are also cases that do not lead exclusively to small values of the double-logarithmic derivatives. Two such scenarios will be studied in the remainder of this appendix.

2. Gaussian correlations

Consider a two-point correlation function of Gaussian shape for a field on a one-dimensional space, i.e.

$$C(r) = C_0 \mathrm{e}^{-\frac{r^2}{2\sigma^2}}. \tag{A4}$$

A field whose statistics are described by such a two-point correlation function can be regarded as a stationary and spatially uncorrelated field convolved with a Gaussian. Calculating the corresponding power spectrum via Fourier transformation yields

$$P_k = \sqrt{2\pi}\sigma C_0 \mathrm{e}^{-\frac{\sigma^2 k^2}{2}}. \tag{A5}$$

This power spectrum drops quickly with increasing k, due to the flatness of the correlation function around $r = 0$. Therefore, both the first and second double-logarithmic derivatives grow unbounded as $k \to \infty$. Thus, by employing a spectral smoothness prior that punishes large values for these derivatives, one prevents in principle the reconstruction of the true power spectrum and supresses the small-scale correlations in the reconstructed field that are in reality more pronounced. In cases in which Gaussian correlations are expected, it might therefore be advisable to choose the strength of the spectral smoothness prior, given by σ_p, k-dependent or choose a different smoothness prior altogether.

3. Triangular correlations

Another case in which the double-logarithmic derivatives of the power spectrum can become divergent is a correlation function with finite support. As an example, we consider a signal on a one-dimensional space with correlations only over the finite distance $2L$, given by

$$C(r) = \begin{cases} C_0 \left(1 - \frac{r}{L}\right) & \text{for } r < L \\ 0 & \text{else} \end{cases}. \tag{A6}$$

This triangular correlation function describes a stationary and spatially uncorrelated field that has been convolved with a top-hat kernel. Fourier transforming it yields the power spectrum

$$P_k = \frac{2C_0}{k^2 L} \left(1 - \cos\left(kL\right)\right). \qquad \text{(A7)}$$

This power spectrum becomes exactly zero at finite k-values. Its logarithm, and therefore also the double-logarithmic derivatives, are divergent at these locations, so not even a k-dependent value of σ_p can ensure the correct reconstruction of the power spectrum in this case. The oscillatory behavior of this power spectrum is a generic feature of signal fields that are correlated only over a finite distance.

Note that both in the case of only locally correlated fields and in the case of Gaussian correlations, the double-logarithmic derivatives can be kept finite by adding a constant floor to the power spectrum, i.e. introducing an additive part to the signal that is spatially uncorrelated. In many cases, the variance of the uncorrelated addition needed to satisfy the spectral smoothness prior, Eq. (24), will be small enough so as not to influence the signal reconstruction significantly. Note that in practice, any field variations are restricted by the finite pixel size and only k-values up to a finite $k_{\max}$ will be considered. In Sec. II C, we investigate the effect that the spectral smoothness prior given in Eq. (24) has in practice on the recunstruction of a signal field exhibiting the two potentially problematic two-point correlations that we discussed here.

In conclusion, using a spectral smoothness prior that punishes large values for the first or second double-logarithmic derivative of the power spectrum can lead to the introduction of spurious small-scale variations in cases in which the true two-point correlation function is flat around $r = 0$ or has only finite support. In the latter case, it can also introduce spurious large-scale correlations. This will, however, in general only be a problem for the reconstruction if some feature that mimics these large-scale correlations is present in the data, i.e. caused by the noise. Another case in which the employment of any spectral smoothness prior is obviously a bad idea is that of a signal that exhibits prominent periodicities. The detection of such spectral lines would only get hindered by the usage of a spectral smoothness prior.

Appendix B: Discretization of the spectral smoothness prior

In our implementation we use discretized values $(k_i)_{i=0,\dots,i_{\max}}$ for the length scales that are represented on the computational grid, or in case of the two-torus, bins thereof. We approximate the integral in the exponent of the spectral smoothness prior, Eq. (24), with a sum according to

$$\int \mathrm{d}(\log k) \left(\frac{\partial^2 \log P_k}{\partial \left(\log k\right)^2}\right)^2 \approx \sum_{i=1}^{i_{\max}-1} \delta_i \left(\Delta p\right)_i^2. \qquad \text{(B1)}$$

Here, we use the abbreviations

$$\delta_i = \frac{\log k_{i+1} - \log k_{i-1}}{2} \qquad \text{(B2)}$$

for integer values of i and

$$\delta_i = \log k_{i+1/2} - \log k_{i-1/2} \qquad \text{(B3)}$$

for half-integer values of i. Note that we have excluded the boundaries at k_0 and $k_{i_{\max}}$ from the sum to avoid numerical problems at these locations.

We approximate the second logarithmic derivative as

$$\left(\Delta p\right)_i = \sum_j \Delta_{i,j} p_j = \frac{\frac{p_{i+1}-p_i}{\delta_{i+1/2}} - \frac{p_i - p_{i-1}}{\delta_{i-1/2}}}{\delta_i}, \qquad \text{(B4)}$$

so that we can represent it as a matrix with the entries

$$\Delta_{i,i} = -\frac{1}{\delta_i}\left(\frac{1}{\delta_{i+1/2}} + \frac{1}{\delta_{i-1/2}}\right), \qquad \text{(B5)}$$

$$\Delta_{i,i\pm1} = \frac{1}{\delta_i \delta_{i\pm1}}, \qquad \text{(B6)}$$

acting on the vector p. All other entries of the matrix Δ are zero.

We can now write the exponent of Eq. (24) as

$$-\frac{1}{2\sigma_p^2} \int \mathrm{d}(\log k) \left(\frac{\partial^2 \log P_k}{\partial \left(\log k\right)^2}\right)^2 = -\frac{1}{2} \sum_{i,j=0} p_i T_{i,j} p_j,$$
$$\text{(B7)}$$

where the matrix T is given by

$$T_{i,j} = \frac{1}{\sigma_p^2} \sum_l \Delta_{l,i} \delta_l \Delta_{l,j}. \qquad \text{(B8)}$$