Signal inference with unknown response: Calibration-uncertainty renormalized estimator

Sebastian Dorn,* Torsten A. Enßlin, Maksim Greiner, Marco Selig, and Vanessa Boehm
Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, D-85748 Garching, Germany
and Ludwigs-Maximilians-Universität München, Geschwister-Scholl-Platz 1, D-80539 München, Germany
(Received 23 October 2014; published 29 January 2015)

The calibration of a measurement device is crucial for every scientific experiment, where a signal has to be inferred from data. We present CURE, the calibration-uncertainty renormalized estimator, to reconstruct a signal and simultaneously the instrument's calibration from the same data without knowing the exact calibration, but its covariance structure. The idea of the CURE method, developed in the framework of information field theory, is to start with an assumed calibration to successively include more and more portions of calibration uncertainty into the signal inference equations and to absorb the resulting corrections into renormalized signal (and calibration) solutions. Thereby, the signal inference and calibration problem turns into a problem of solving a single system of ordinary differential equations and can be identified with common resummation techniques used in field theories. We verify the CURE method by applying it to a simplistic toy example and compare it against existent self-calibration schemes, Wiener filter solutions, and Markov chain Monte Carlo sampling. We conclude that the method is able to keep up in accuracy with the best self-calibration methods and serves as a noniterative alternative to them.

DOI: 10.1103/PhysRevE.91.013311 PACS number(s): 02.70.−c, 02.50.−r, 05.10.−a, 07.05.Kf

I. INTRODUCTION

A. Motivation

Data analysis is the link between theory and experiment, wherein a signal has to be inferred from measured data. For this purpose the transformation of a signal to data, the measurement response, has to be understood precisely. The reconstruction of this response is called calibration.

In the simplest case of a time-independent instrument response, the calibration can be determined by measuring an *a priori* well-known signal in a regime with negligible noise level. This is commonly called external calibration. However, the assumption of time independency cannot be accepted in the majority of cases. Of course the time dependency caused by, e.g., environmental factors, periodicities, and systematics, or the signal itself, can be estimated with utmost effort. The resulting calibration, however, has still to be extrapolated into future time, where the real measurement will be performed and where these influences will not be known exactly. What might be known, however, is their statistics. The resulting uncertainty consequently affects the signal reconstruction and has to be taken into account.

There are methods which improve the calibration by iteratively calibrating on a signal reconstruction and then improving the reconstruction using the new calibration. Such self-calibration (selfcal) schemes are widely in usage. They can, however, be prone to systematic biases since signal and calibration are partly degenerate, i.e., a feature in the data could be caused by either of them and it is not guaranteed that the selfcal scheme makes the correct choice automatically.

An improved selfcal scheme, which takes signal uncertainties in the calibration step into account, was presented in Ref. [1]. Since this new selfcal is also an approximative solution to the complex inference problem, we ask if there is room for further improvement using information field theory (IFT) [2]. To this end we develop a calibration-uncertainty renormalized estimator (CURE) for a signal, which incorporates calibration uncertainties successively in a so-called renormalization flow equation. In comparison to existent approaches this method is noniterative. For a review and discussion of previous work on existent calibration methods, we point to Refs. [1,3].

B. Structure of the work

The remainder of this work is organized as follows. In Sec. II we review the basics of the free and interacting IFTs with focus on the latter. Section III represents the main part of the paper, where the calibration problem is introduced and the CURE method is derived. The basic ideas as well as the main formulas of alternative selfcal schemes are also presented within this section. In Sec. IV the performance of several signal reconstruction methods is studied within a numerical toy example. Results are summarized in Sec. V.

II. INFORMATION FIELD THEORY

To follow the derivation of an estimator with renormalized calibration uncertainty in the framework of IFT one has to be familiar with the concepts of interacting IFT (see in particular Secs. II B and II C). Thus, a brief review might be helpful, but can be skipped by an advanced reader. For this purpose we basically follow Refs. [2,4], where a more detailed description of IFT can be found.

A. Basic formalism and free theory

Typically, a signal has to be inferred from data with the challenging question: how can this be done in an optimal[1] way? To reasonably answer this question we first have to agree on a particular data model.

Within this work we assume that the data can be expressed by a discrete data-tuple, $d = (d_1, \ldots, d_m)^T \in \mathbb{R}^m$, $m \in \mathbb{N}$,

*sdorn@mpa-garching.mpg.de

[1]Optimal with respect to, e.g., minimizing the $\mathcal{L}^2$ error.

which is related to a signal s by

$$d = Rs + n, \tag{1}$$

where R is a linear response operation acting on the signal, and $n = (n_1, \ldots, n_m)^T \in \mathbb{R}^m$ denotes some measurement noise. In contrast to data and noise, the signal $s \equiv s(x)$, $x \in \mathcal{U}$ is considered to be a continuous quantity over some Riemannian manifold $\mathcal{U}$, i.e., a physical (scalar) field. The linearity of the signal response, which transforms the continuous signal into data space, is valid for many physical measurements, e.g., observations of the cosmic microwave background and large-scale structure in astronomy (cosmology), spectroscopy in different fields of physics, or medical imaging.

We further assume the signal and noise to be uncorrelated, $\mathcal{P}(s,n) = \mathcal{P}(s)\mathcal{P}(n)$, and primarily Gaussian, i.e., $\mathcal{P}(s) = \mathcal{G}(s,S)$ and $\mathcal{P}(n) = \mathcal{G}(n,N)$ with related covariances $S = \langle ss^\dagger \rangle_{(s|S)}$ and $N = \langle nn^\dagger \rangle_{(n|N)}$, respectively. Here, we implicitly introduced the notation

$$\mathcal{G}(a,A) \equiv \frac{1}{\sqrt{|2\pi A|}} \exp\left(-\frac{1}{2} a^\dagger A^{-1} a \right) \quad \text{and}$$
$$\langle \cdot \rangle_{(a|A)} \equiv \int \mathcal{D}a \cdot \mathcal{P}(a|A), \tag{2}$$

where $\dagger$ denotes a transposition and $*$ complex conjugation. The appropriate inner product of two fields $\{a,b\}$ is defined by $a^\dagger b \equiv \int_{\mathcal{U}} d^{\dim \mathcal{U}} x a^*(x) b(x)$. If the conditions described above (known linear response, Gaussian signal, and noise with known covariances) are met, we term the theory a *free theory*.

It is often convenient and common to focus on logarithmic probabilities by relating Bayes' theorem [5] to statistical physics,

$$\mathcal{P}(s|d) = \frac{\mathcal{P}(s,d)}{\mathcal{P}(d)} \equiv \frac{1}{\mathcal{Z}} \exp[-\mathcal{H}(s,d)]. \tag{3}$$

Here, we introduced the information Hamiltonian

$$\mathcal{H}(s,d) \equiv -\ln[\mathcal{P}(s,d)] \tag{4}$$

and the partition function

$$\mathcal{Z}(d) \equiv \mathcal{P}(d) = \int \mathcal{D}s \exp\left[-\mathcal{H}(s,d)\right]. \tag{5}$$

Still considering the above free theory, we find

$$\mathcal{H}(s,d) = \mathcal{H}_0 - j^\dagger s + \tfrac{1}{2} s^\dagger D^{-1} s \quad \text{and}$$
$$\mathcal{Z}(d) = \sqrt{|2\pi D|} \exp\left(\tfrac{1}{2} j^\dagger D j - \mathcal{H}_0\right), \tag{6}$$

with the abbreviations

$$\mathcal{H}_0 = \tfrac{1}{2} \ln|2\pi N| + \tfrac{1}{2} \ln|2\pi S| + \tfrac{1}{2} d^\dagger N^{-1} d,$$
$$D^{-1} = S^{-1} + R^\dagger N^{-1} R, \quad \text{and}$$
$$j^\dagger = d^\dagger N^{-1} R, \tag{7}$$

where the so-called information propagator D and the information source j have been introduced. $|\cdot|$ denotes the determinant.

To exploit the whole machinery of statistical physics we additionally include a moment-generating term $J^\dagger s$ into the partition function,

$$\mathcal{Z}(d,J) = \int \mathcal{D}s \exp\left[-\mathcal{H}(s,d) + J^\dagger s\right]. \tag{8}$$

The last definition permits the connected correlation functions (the cumulants) of a probability density function (PDF) to be expressed via functional derivatives [2],

$$\langle s(x_1) \cdots s(x_n) \rangle^c_{(s|d)} \equiv \left. \frac{\delta^n \ln[\mathcal{Z}(d,J)]}{\delta J(x_1) \cdots \delta J(x_n)} \right|_{J=0}. \tag{9}$$

Since we consider a Gaussian signal, its mean is equivalent to the well-known Wiener filter [6] solution,

$$\langle s \rangle_{(s|d)} = Dj \equiv m_{\mathrm{w}}. \tag{10}$$

Its two-point correlation function describes the uncertainty of the reconstruction, $\langle ss^\dagger \rangle^c_{(s|d)} = \langle (s - m_{\mathrm{w}})(s - m_{\mathrm{w}})^\dagger \rangle_{(s|d)} = D$, and all cumulants with $n > 2$ vanish. Therefore, the posterior is Gaussian and given by

$$\mathcal{P}(s|d) = \mathcal{G}(s - m_{\mathrm{w}}, D). \tag{11}$$

B. *n*th-order perturbation theory

Within the free theory we required the noise and in particular the signal to be Gaussian. However, this requirement cannot be met in some cases, e.g., in the cases where noise or response is signal dependent, or simply for a nonlinear signal field. In the framework of IFT these scenarios can often[2] be described by a Taylor-expanded Hamiltonian [2] composed of a free part, $\mathcal{H}_{\mathrm{free}}$ [Eq. (6)] and a so-called interacting part $\mathcal{H}_{\mathrm{int}}$,

$$\mathcal{H} = \mathcal{H}_{\mathrm{free}} + \underbrace{\sum_{n=0}^{\infty} \frac{1}{n!} \Lambda^{(n)}[s^{(n)}]}_{\equiv \mathcal{H}_{\mathrm{int}}}, \tag{12}$$

where the deviation from Gaussianity is encoded in the anharmonic terms [$n > 2$]. The term $\Lambda^{(n)}[s^{(n)}]$ denotes a complete, fully symmetric,[3] contraction between the rank-n tensor $\Lambda^{(n)}$ and the n fields $s^{(n)} = (s^1, \ldots, s^n)$. If a decent estimate m_0 is known, one should Taylor-expand the Hamiltonian around this reference field m_0 in terms of residuals $\phi \equiv s - m_0$. An estimate that works well is, for instance, the Wiener filter solution of the free theory, Eq. (10). Using this reference field expansion often permits earlier truncation of the Taylor expansion, since the anharmonic terms become smaller.

Analogously to the free theory, we define the partition function

$$\begin{aligned}
\mathcal{Z}(d,J) &= \int \mathcal{D}s \exp\left[-\mathcal{H}(s,d) + J^\dagger s\right] \\
&= \int \mathcal{D}s \exp[-\mathcal{H}_{\mathrm{int}}] \exp\left[-\mathcal{H}_{\mathrm{free}} + J^\dagger s\right] \\
&= \exp\left(-\mathcal{H}_{\mathrm{int}}\left[\frac{\delta}{\delta J}\right]\right) \int \mathcal{D}s \exp\left[-\mathcal{H}_{\mathrm{free}} + J^\dagger s\right] \\
&\equiv \exp\left(-\mathcal{H}_{\mathrm{int}}\left[\frac{\delta}{\delta J}\right]\right) \mathcal{Z}_{\mathrm{free}} \\
&= \left(1 - \mathcal{H}_{\mathrm{int}}\left[\frac{\delta}{\delta J}\right] + \frac{1}{2!}\mathcal{H}^2_{\mathrm{int}}\left[\frac{\delta}{\delta J}\right] - \cdots\right) \mathcal{Z}_{\mathrm{free}}.
\end{aligned} \tag{13}$$

[2] See Sec. II C for cases in which such a treatment is not sufficient.

[3] $\Lambda^{(n)} \equiv (1/n!) \sum_\pi \Lambda^{(n)}_{\pi(x_1,\ldots,x_n)}$, with π representing every permutation of $\{1, \ldots, n\}$.

In principle, Eqs. (9) and (13) enable the calculation of all correlation functions of a PDF perturbatively. These calculations, however, are very difficult and lengthy. Fortunately, there exists a well-known diagrammatic treatment in analogy to quantum field theory and thermal field theory [2]. For example, including the first two correction terms, the signal mean m is given by

$$= D_{xy}\left(j_y - \frac{1}{2}\Lambda^{(3)}_{yzv}D_{zv} - \frac{1}{2}\Lambda^{(3)}_{yzv}(Dj)_z(Dj)_v\right) + \cdots, \tag{14}$$

where the ordering of diagrams corresponds to those of the equations and the ellipses $(\cdots)$ representing the residual Feynman series of correction terms. The external dots $(\bullet)$ represent source terms, the internal dots vertices (the tensors $\Lambda^{(n)}$), and the lines (—) propagator terms, respectively. Repeated indices are to be integrated over.

The Feynman rules used in this work, which are necessary to switch between the mathematical expressions and the corresponding diagrams, can be found in Appendix A.

C. Uncertainty renormalization

1. Motivation

The approach of perturbative diagrammatic expansion is supposed to work well if the Hamiltonian is dominated by linear and quadratic terms. That in turn means that the tensors $\Lambda^{(n)}$ describing the deviation from Gaussianity are sufficiently small for the Feynman series to converge. This is, however, not always the case, e.g., within the calibration problem where the signal response cannot be known exactly due to some potential time dependencies or uncontrolled systematics. This calibration uncertainty can lead to large, nonvanishing terms $\Lambda^{(n)}$ as we show in Sec. III A of this paper.

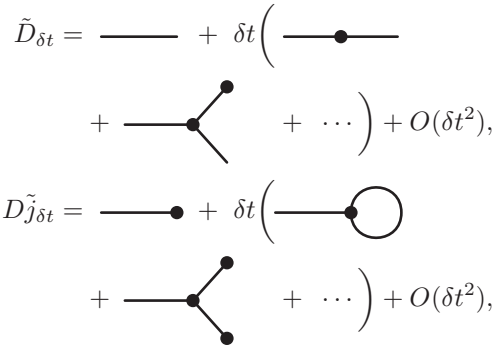
Following the concept of Ref. [4], we can circumvent this problem by including successively more and more small portions of, e.g., calibration uncertainty into a signal inference equation. The basic idea is to include only a sufficiently small amount of uncertainty per step to ensure that the non-Gaussian (interaction) terms are weak. Finally, this process results in a renormalized propagator $\tilde{D}$ and information source $\tilde{j}$. This process is called uncertainty renormalization [4].

2. Concept

For reasons of clarity and comprehensibility we skip the most general derivation and justification of uncertainty renormalization, which can be found in Ref. [4], and focus more on the pragmatic procedure thereof. In the following we consider the Taylor-expanded, effective Hamiltonian to be of the form of Eq. (12). To suppress the strength of the non-Gaussian contributions we include a so-called expansion parameter $\delta t \ll 1$, into the Hamiltonian,

$$\mathcal{H} = \mathcal{H}_{\text{free}} + \delta t \sum_{n=0}^{\infty} \frac{1}{n!}\Lambda^{(n)}[s^{(n)}], \tag{15}$$

and concentrate on this new Hamiltonian for a moment. For an appropriately small δt the interaction terms become sufficiently small and the diagrammatic expansion of Sec. II B is justified again. Hence, by including the first correction terms into the propagator, $D \to \tilde{D}_{\delta t}$, and into the information source, $j \to \tilde{j}_{\delta t}$, we obtain



$$+ O(\delta t^2), \tag{16}$$

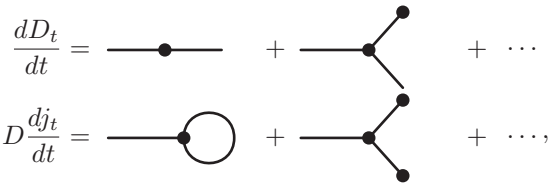
where the ellipses $(\cdots)$ represent all diagrams of order $O(\delta t)$, i.e., all possible one-vertex diagrams. In this way, $t \in [0,1]$ can be identified with a pseudotime, which measures the accumulated uncertainty correction to the information propagator and source, and the expansion parameter δt represents the time step in which D and j are increased from their intermediate values D_t and j_t to their one-step-renormalized (but not final) values $D_{t+\delta t}$ and $j_{t+\delta t}$, i.e.,

$$D_t \to D_{t+\delta t} \quad \text{and} \quad j_t \to j_{t+\delta t}. \tag{17}$$

We want to emphasize that δt cannot simply be set to unity to obtain the fully renormalized propagator $\tilde{D}$, because this step would violate the justification of our perturbative expansion (see Sec. II C 1). However, a single step of this analytical *resummation* can be infinitesimally small, permitting for the formal definition of the derivatives [4]

$$\frac{dD_t}{dt} \equiv \lim_{\delta t \to 0} \frac{D_{t+\delta t} - D_t}{\delta t} \quad \text{and}$$
$$\frac{dj_t}{dt} \equiv \lim_{\delta t \to 0} \frac{j_{t+\delta t} - j_t}{\delta t}, \tag{18}$$

whereby the *renormalization flow equations* can be formulated,



$$D\frac{dj_t}{dt} = \quad + \quad + \cdots, \tag{19}$$

which is a system of coupled differential equations for operators with boundary values $D_{t=0} = D$ and $j_{t=0} = j$. By solving these equations one obtains the fully renormalized quantities $\tilde{D} = D_{t=1}$, $\tilde{j} = j_{t=1}$, and the renormalized Wiener filter formula

$$\tilde{m} = \tilde{D}\tilde{j}. \tag{20}$$

This means that by solving Eq. (19) we finally calculate a Gaussian approximation to the correct posterior mean of s, $P(s|d) \approx \mathcal{G}(s - \tilde{m}, \tilde{D})$.

III. SELF-CALIBRATION

Now we address the calibration problem, i.e., how to infer a physical signal field given a data set without precise knowledge of the signal response. We consider the case in which an external calibration is not possible (see Sec. I). Thus, the instrument has to be self-calibrated during the measurement process. If we had absolutely no information about the signal response (how a measurement device transforms the signal into data) there would be absolutely no chance to infer the signal appropriately. However, if we have some information about the statistics of the response, e.g., the two-point correlation function, this task becomes solvable. For this purpose we introduce the CURE method in the framework of IFT (Sec. III A) and review already existing methods (Sec. III B) to compare it against.

The aim is to calculate an optimal[4] estimator for the signal (or in general the moments $\langle s \cdots s \rangle_{(s|d)}$) given the data without exact information of the calibration. A way to approach this challenge is to consider the unknown calibration as a nuisance parameter, i.e., to marginalize over the calibration when calculating the signal posterior,

$$\mathcal{P}(s|d) = \int \mathcal{D}\gamma \mathcal{P}(s,\gamma|d) = \underbrace{\int \mathcal{D}\gamma \mathcal{P}(d,\gamma|s)}_{\mathcal{P}(d|s)} \frac{\mathcal{P}(s)}{\mathcal{P}(d)}, \quad (21)$$

which involves the calculation of the calibration marginalized likelihood. To do so, we assume the response to be a linear function in the calibration coefficients γ_a with Gaussian statistics, i.e., $R^\gamma \approx R^0 + \sum_a \gamma_a R^a$. The assumption of Gaussianity is appropriate as long as we have *a priori* no information about higher moments of γ, $\langle \gamma_1 \cdots \gamma_n \rangle_{(\gamma)}$ with $n > 2$. The linearity can be considered as a first-order approximation around $\gamma_0 = 0$ in γ,

$$\begin{aligned} R^\gamma &= R(\gamma_0) + \left.\frac{\partial R(\gamma)}{\partial \gamma_a}\right|_{\gamma=\gamma_0} (\gamma - \gamma_0) + O(\gamma^2) \\ &= R^0 + \sum_a \gamma_a R^a + O(\gamma^2). \end{aligned} \quad (22)$$

Under these assumptions one obtains [1,7]

$$\begin{aligned} \mathcal{P}(d|s) &= \int \mathcal{D}\gamma \mathcal{P}(d|s,\gamma)\mathcal{P}(\gamma) \\ &= \int \mathcal{D}\gamma \mathcal{G}\left[d - \left(R^0 + \sum_a \gamma_a R^a\right) s, N\right] \mathcal{G}(\gamma,\Gamma) \\ &= \mathcal{G}\left(d - R^0 s, N + \sum_{ab} \Gamma_{ab} R^a s s^\dagger R^{b\dagger}\right). \end{aligned} \quad (23)$$

The data variance of this Gaussian likelihood, Eq. (23), depends on the correlation structure of the calibration, $\Gamma = \langle \gamma\gamma^\dagger \rangle_{(\gamma|\Gamma)}$, as well as on the signal s. This, in turn, results in a non-Gaussian posterior $\mathcal{P}(s|d) \propto \mathcal{P}(d|s)\mathcal{P}(s)$, such that calculations of moments cannot be done analytically anymore. In principle one can adapt posterior sampling techniques like Markov chain Monte Carlo (MCMC) methods to calculate, e.g., the posterior mean, m_{MCMC}. These approaches, however, are usually very expensive, which increases the attractivity of developing (semi)analytical methods.

A. Calibration-uncertainty renormalized estimator

Now, we apply the concept of uncertainty renormalization to the selfcal problem. According to Sec. II C we introduce an expansion parameter $\delta t \ll 1$ in the ansatz:

$$\mathcal{P}(s|d) \propto \mathcal{G}\left(d - R^0 s, N + \delta t \sum_{ab} \Gamma_{ab} R^a s s^\dagger R^{b\dagger}\right) \mathcal{P}(s). \quad (24)$$

To simplify the notation we define an auxiliary parameter $\Xi \equiv \sum_{ab} \Gamma_{ab} R^a s s^\dagger R^{b\dagger}$ and assume a Gaussian signal prior $\mathcal{P}(s) = \mathcal{G}(s - s_0, S)$, with the *a priori* mean $s_0 \equiv \langle s \rangle_{(s)}$.

The Hamiltonian becomes

$$\begin{aligned} \mathcal{H}(d,s) &= -\ln \mathcal{P}(d,s) \\ &= -\ln[\mathcal{G}(d - R^0 s, N + \delta t\, \Xi)\mathcal{G}(s - s_0, S)] \\ &= \tfrac{1}{2} \ln |2\pi S| + \tfrac{1}{2} \ln |2\pi (N + \delta t\, \Xi)| \\ &\quad + \tfrac{1}{2}(d - R^0 s)^\dagger (N + \delta t\, \Xi)^{-1} (d - R^0 s) \\ &\quad + \tfrac{1}{2}(s - s_0)^\dagger S^{-1} (s - s_0). \end{aligned} \quad (25)$$

We can use the fact that the expansion parameter δt is small, i.e., $\delta t\, \Xi \ll N$ (spectrally[5]), whereby the approximations

$$\begin{aligned} \ln |2\pi (N + \delta t\, \Xi)| &\approx \ln |2\pi N| + \mathrm{tr}(\delta t\, \Xi N^{-1}) \quad \text{and} \\ (N + \delta t\, \Xi)^{-1} &\approx N^{-1} - N^{-1} \delta t\, \Xi N^{-1} \end{aligned} \quad (26)$$

can be made. Use of Eqs. (25) and (26) yields

$$\mathcal{H}(d,s) = \mathcal{H}^{\mathrm{free}} + \delta t \sum_{n=2}^{4} \frac{1}{n!} \lambda^{(n)}[s^{(n)}] \quad (27)$$

with

$$\begin{aligned} \mathcal{H}^{\mathrm{free}} &= \mathcal{H}_0 + \frac{1}{2} s^\dagger D^{-1} s - j^\dagger s, \\ \lambda^{(2)}[s,s] &= \sum_{ab} \Gamma^{ab} (s^\dagger M^{ba} s - j^{a\dagger} s s^\dagger j^b) + 1 \text{ perm}, \\ \lambda^{(3)}[s,s,s] &= \sum_{ab} \Gamma^{ab} \left(\frac{1}{2} j^{a\dagger} s s^\dagger M^{b0} s + \text{c.c.}\right) + 5 \text{ perm}, \\ \lambda^{(4)}[s,s,s,s] &= \sum_{ab} \Gamma^{ab} \left(-\frac{1}{2} s^\dagger M^{0a} s s^\dagger M^{b0} s\right) + 23 \text{ perm}, \end{aligned} \quad (28)$$

with permutations (perm) with respect to s and the abbreviations

$$\begin{aligned} \mathcal{H}_0 &= \tfrac{1}{2} \ln |2\pi N| + \tfrac{1}{2} \ln |2\pi S| + \tfrac{1}{2} d^\dagger N^{-1} d + \tfrac{1}{2} s_0^\dagger S^{-1} s_0, \\ D^{-1} &= (S^{-1} + R^{0\dagger} N^{-1} R^0), \\ j^\dagger &= d^\dagger N^{-1} R^0 + s_0^\dagger S^{-1}, \end{aligned}$$

[4] Optimal in the sense of minimizing the $\mathcal{L}^2$ error.

[5] Meaning that $\xi^\dagger \delta t\, \Xi \xi \ll \xi^\dagger N \xi \;\forall\; \xi \in \mathbb{R}^m \backslash 0$.

$$M^{ab} = R^{a\dagger} N^{-1} R^b,$$
$$j^{a\dagger} = d^\dagger N^{-1} R^a. \tag{29}$$

Terms higher than fourth order in the signal are dropped by making the approximation of Eq. (26).

1. Zero-point expansion

Since the information Hamiltonian [Eqs. (27), (28), and (29)] has the structure of Eq. (15), we can start to derive the renormalization flow equations. First, we consider (also for pedagogical reasons) the special case, in which the *a priori* signal mean is zero but the signal two-point statistic is known, i.e., we use a zero-centered, Gaussian prior $\mathcal{P}(s) = \mathcal{G}(s,S)$.

Following Sec. II C, the interaction terms of Eq. (27) [Eq. (28)] can be absorbed in a so-called renormalized information propagator $\tilde{D}_{\delta t}$ and information source $\tilde{j}_{\delta t}$ of order δt. Including this (first) correction these quantities read

$$\begin{aligned}
\left(\tilde{D}_{\delta t}\right)_{xy} &= D_{xy} + \delta t \left(-D_{xz}\lambda^{(2)}_{zv} D_{vy} - D_{xz}\lambda^{(3)}_{zvu}(Dj)_v D_{uy} - \frac{1}{2} D_{xz}\lambda^{(4)}_{zvur} D_{vu} D_{ry} - \frac{1}{2} D_{xz}\lambda^{(4)}_{zvur}(Dj)_v (Dj)_u D_{ry} \right) \\
&\quad + O\left(\delta t^2\right) \\
&= \text{—} + \delta t \Big(\text{—•—} + \text{Y} + \text{—◯} + \text{+} \Big) + O\left(\delta t^2\right), \\
D_{xy}\left(\tilde{j}_{\delta t}\right)_y &= D_{xy} \Big[j_y + \delta t \Big(-\frac{1}{2}\lambda^{(3)}_{yzv} D_{zv} - \lambda^{(2)}_{yz}(Dj)_z - \frac{1}{2}\lambda^{(3)}_{yzv}(Dj)_z (Dj)_v - \frac{1}{2}\lambda^{(4)}_{yzvu} D_{zv}(Dj)_u \\
&\quad - \frac{1}{3!}\lambda^{(4)}_{yzvu}(Dj)_z (Dj)_v (Dj)_u \Big) \Big] + O\left(\delta t^2\right) \\
&= \text{—•} + \delta t \Big(\text{•◯} + \text{—•—} + \text{Y} + \text{•◯} + \text{+} \Big) \\
&\quad + O\left(\delta t^2\right).
\end{aligned} \tag{30}$$

Just as a reminder, the vertices (internal dots) are multiplied by δt while the source terms (external dots) are independent of δt. In the diagrammatic expansions Eq. (30), we place δt outside the parentheses to underline this dependency. Therefore, to include all corrections up to order δt, we have to include all possible one-vertex diagrams. It is crucial to realize that δt cannot simply be set to 1 in order to obtain the fully renormalized propagator $\tilde{D}$, because this step would violate Eq. (26). Apart from this it might also break down the perturbative expansion. However, instead of setting $\delta t = 1$ we can interpret $t \in [0,1]$ as a pseudotime, which measures the accumulated correction to the information propagator and source (see Sec. II C), $D_{t+\delta t}$ and $j_{t+\delta t}$. Thereby we can formulate the renormalization flow equations,

$$\begin{aligned}
\frac{dD_t}{dt} &= \lim_{\delta t \to 0} \frac{D_{t+\delta t} - D_t}{\delta t} \\
&= \text{—•—} + \text{Y} \\
&\quad + \text{—◯} + \text{+} \quad , \\
D\frac{dj_t}{dt} &= D\left(\lim_{\delta t \to 0} \frac{j_{t+\delta t} - j_t}{\delta t} \right) \\
&= \text{•◯} + \text{—•—} + \text{Y} \\
&\quad + \text{—◯} + \text{+} \quad ,
\end{aligned} \tag{31}$$

which is a system of coupled differential equations for operators with boundary values $D_{t=0} = D$ and $j_{t=0} = j$. By solving these equations one obtains the fully renormalized quantities $\tilde{D} = D_{t=1}$, $\tilde{j} = j_{t=1}$, and the renormalized Wiener filter formula

$$\tilde{m} = \tilde{D}\tilde{j}. \tag{32}$$

However, instead of solving the coupled differential equations of Eq. (31) we could also solve the system where dD_t/dt is replaced by an equivalently valid equation for dD_t^{-1}/dt leading to the new differential system

$$\begin{aligned}\frac{dD_{t,xy}^{-1}}{dt} &= \lambda_{xy}^{(2)} + \lambda_{xyz}^{(3)}(D_t j_t)_z + \frac{1}{2}\lambda_{xury}^{(4)} D_{t,ur} \\ &\quad + \frac{1}{2}\lambda_{xvuy}^{(4)}(D_t j_t)_v (D_t j_t)_u, \\ \frac{dj_{t,y}}{dt} &= -\frac{1}{2}\lambda_{yzv}^{(3)} D_{t,zv} - \frac{1}{2}\lambda_{yzv}^{(3)}(D_t j_t)_z (D_t j_t)_v \\ &\quad - \lambda_{yz}^{(2)}(D_t j_t)_z - \frac{1}{2}\lambda_{yzvu}^{(4)} D_{t,zv}(D_t j_t)_u \\ &\quad - \frac{1}{3!}\lambda_{yzvu}^{(4)}(D_t j_t)_z (D_t j_t)_v (D_t j_t)_u. \end{aligned} \tag{33}$$

Solving these equations might simplify the numerical effort in some cases. Afterwards we invert $D_{t=1}^{-1} \equiv \tilde{D}^{-1}$ to finally solve Eq. (32).

2. *Reference field expansion*

There is also the option to introduce a residual field $\phi = s - m_0$ with respect to a reference field, e.g., $m_0 = Dj^0$, the Wiener filter solution without information of the proper calibration [Eq. (10)]. By deriving a Hamiltonian of ϕ the perturbative expansion gets more exact while the non-Gaussian terms become smaller. The Hamiltonian then reads

$$\mathcal{H}(d,\phi) = \mathcal{H}_0' + \frac{1}{2}\phi^\dagger D^{-1}\phi + \delta t \sum_{n=1}^{4} \frac{1}{n!}\Lambda^{(n)}[\phi^{(n)}], \tag{34}$$

where $\mathcal{H}_0'$ includes all ϕ-independent terms[6] and $\Lambda^{(n)}$ denotes the new (vertex) tensor. Due to the fact that now already the source term is of $O(\delta t)$, the diagrammatic expansion up to order δt reduces to a sum of Feynman diagrams containing only a single source and single vertex term, given by

$$\begin{aligned}\left(\tilde{D}_{\delta t}\right)_{xy} &= D_{xy} \\ &+ \delta t\left(-D_{xz}\Lambda_{zv}^{(2)} D_{vy} - \frac{1}{2}D_{xz}\Lambda_{zvur}^{(4)} D_{vu} D_{ry}\right) \\ &= \quad\text{—}\quad + \delta t\Big(\quad\text{—}\bullet\text{—}\quad + \quad\text{◯}\quad\Big), \\ D_{xy}\left(\tilde{j}_{\delta t}\right)_y &= D_{xy}\left[\delta t\left(-\Lambda_y^{(1)} - \frac{1}{2}\Lambda_{yzv}^{(3)} D_{zv}\right)\right] \\ &= \delta t\Big(\quad\text{—}\bullet\text{—}\quad + \quad\text{—◯}\quad\Big). \end{aligned} \tag{35}$$

After restoring the original signal s by replacing the source term $j_{\delta t} \to j_{\delta t} + D_{\delta t}^{-1} m_t$, $m_t \equiv D_t j_t$ [1], this leads in analogy to the previous section to the renormalization flow equations,

$$\begin{aligned}\frac{dD_{t,xy}}{dt} &= -D_{xz}\Lambda_{zv}^{(2)} D_{vy} - \frac{1}{2}D_{xz}\Lambda_{zvur}^{(4)} D_{vu} D_{ry}, \\ &\text{or, alternatively,} \\ \frac{dD_{t,xy}^{-1}}{dt} &= \Lambda_{xy}^{(2)} + \frac{1}{2}\Lambda_{xury}^{(4)} D_{t,ur}, \text{ and} \\ \frac{dj_{t,y}}{dt} &= -\Lambda_y^{(1)} + \Lambda_{yz}^{(2)}(D_t j_t)_z \\ &\quad - \frac{1}{2}\Lambda_{yzv}^{(3)} D_{t,zv} + \frac{1}{2}\Lambda_{yzvu}^{(4)} D_{t,zv}(D_t j_t)_u, \end{aligned} \tag{36}$$

with boundaries $j_{t=0} = j^0$ and $D_{t=0}^{-1} = D^{-1}$. Note that the positive terms in the differential equation of j_t arise from the restoration of the original signal.

Further note that the simplicity gained in the diagrammatic expansion has turned into a higher complexity of the vertex structure. The explicit structure of these vertices can be found in Appendix B. These are also implemented for our numerical example; see Sec. IV and Fig. 1. The effect of the resummation process (involving absolute calibration measurements; see Sec. IV) on the information propagator is illustrated by Fig. 2.

3. *Approach optimization*

Until now, the vertex tensors $\Lambda^{(n)}$ were pseudotime independent [see, for instance, Eq. (B2)]. However, the CURE approach can in principle be improved if we replace the residual field $\phi = s - m_0$ after every time step by $\tilde{\phi}_t = s - m_t = s - D_t j_t$. In this way the support point of the expansion is always chosen optimally so that the first term of $\Lambda^{(1)}$ does still vanish [see Eq. (B2)] and the definition of time derivatives, Eq. (31), remains valid.

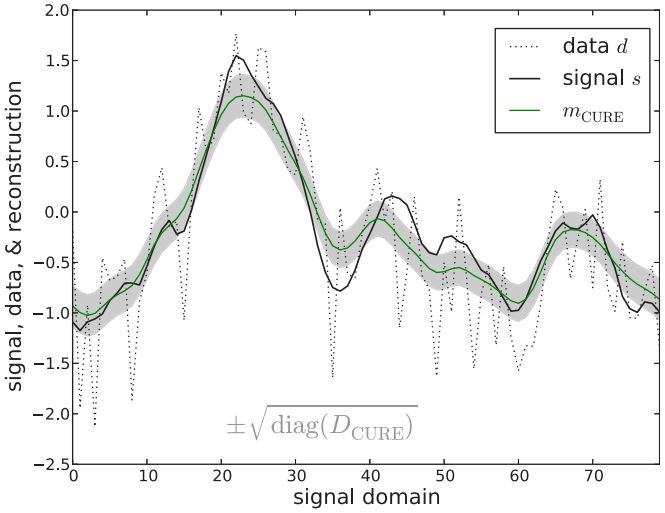


FIG. 1. (Color online) Signal, data, and signal reconstruction (considering unknown calibration) with related 1σ uncertainty according to Eq. (36) for the numerical example described in Sec. IV. For comparison to other methods see Fig. 3. Calibration and its reconstructions are shown in Fig. 4.

[6]Note that among the ϕ-independent terms of $\mathcal{H}_0'$ are terms, collected in $\Lambda^{(0)}$, that depend on δt. These terms, however, only shift the Hamiltonian by a constant value but do not influence its shape or structure.

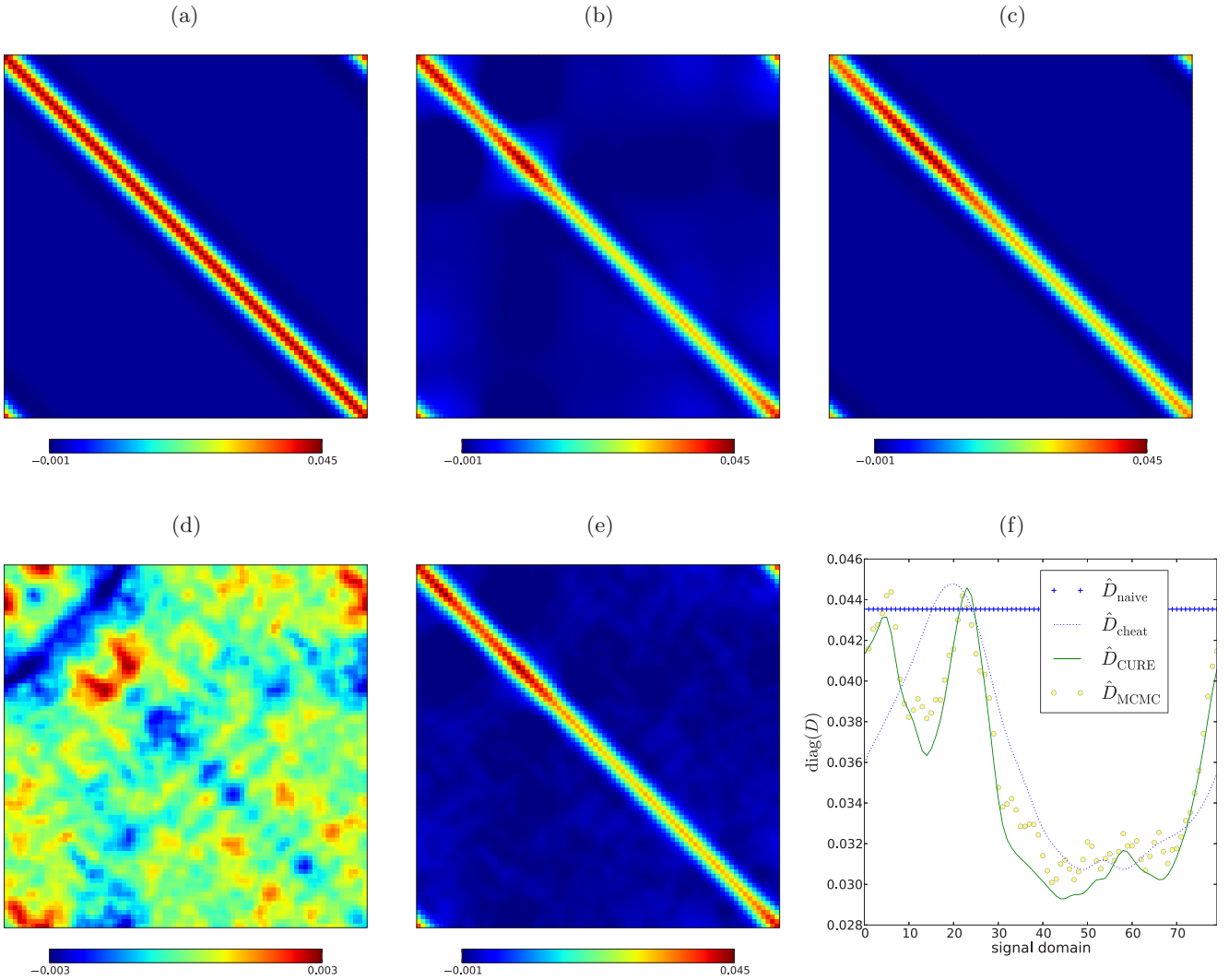


FIG. 2. (Color online) Explicit structure of propagator operators for the realization shown in Figs. 1, 3, and 4. (a)–(c),(e) refer to the propagators D_{method} with the method naive (a),CURE (b),cheat (c), and MCMC (e) according to Eq. (10) with unknown calibration set to zero, Eq. (36), and Eq. (10) with known calibration, respectively. Upper panels and lower middle panel: The renormalized propagator exhibits the same diagonal structure as the MCMC propagator. Lower panels (left, right): Comparison of the renormalized propagator to the MCMC result [CURE vs MCMC, (d)] and explicit structures of the propagator diagonals [denoted by $\hat{D}_{\text{method}}$, (f)]. Emerging from the process of resummation (involving absolute calibration measurements; see Sec. IV), Eq. (36), the renormalized propagator obtains a nondiagonal structure due to the complex, nonlocal vertex structure of the non-Gaussian contributions to the Hamiltonian; see, in particular, Eq. (B2).

B. Self-calibration schemes

To compare the derived CURE method not only to the Wiener filter solution without information about calibration (which is the starting value of CURE) but also to two other iterative self-calibration (selfcal) schemes we review the basic ideas of the latter briefly. A full description of the following methods can be found in Ref. [1]. The response is still considered to be linear, $R^{\gamma} = R^0 + \sum_a \gamma_a R^a$.

1. Classical selfcal

Classical selfcal is an iterative method, alternately inferring the signal while assuming the calibration to be known and vice versa until a fixpoint is reached. The respectively inferred quantities $s^\star$ and $\gamma^\star$ are often maximum *a posterori* (MAP) estimators. This procedure of simultaneously estimating s and γ can be identified with searching for the maximum of the joint posterior $\mathcal{P}(\gamma,s|d)$, or equally for the minimum of the joint information Hamiltonian [1] $\mathcal{H}(d,\gamma,s) = -\ln[\mathcal{P}(d,\gamma,s)]$, given by

$$\left.\frac{\partial\mathcal{H}(d,\gamma,s)}{\partial\gamma_a}\right|_{\gamma=\gamma^\star} = 0 \quad \text{and} \quad \left.\frac{\partial\mathcal{H}(d,\gamma,s)}{\partial s}\right|_{s=s^\star} = 0. \tag{37}$$

The resulting equations [Eq. (39) with $T = 0$] must be iterated until convergence.

2. New selfcal

The new selfcal method is based on the above-described idea of classical selfcal. However, in marked contrast to the latter, new selfcal uses the signal marginalized posterior to infer the calibration, and determines a signal estimate

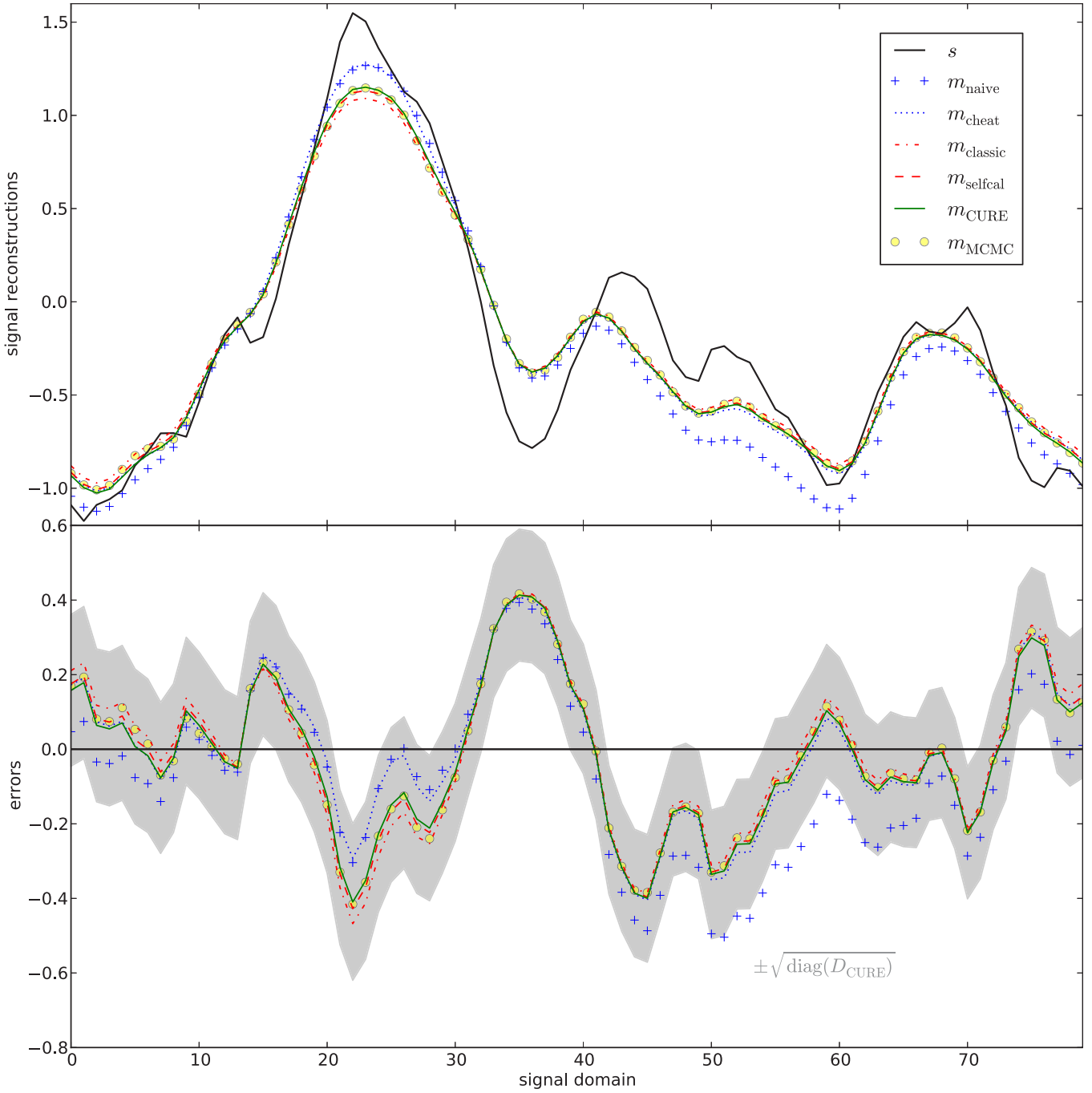


FIG. 3. (Color online) Signal reconstructions and related errors of different approaches. The following terminology is used: Naive, Wiener filter with unknown calibration set to zero; classic, classical selfcal [Eq. (43), $T = 0$]; CURE [Eqs. (36) and (B2)]; selfcal, new selfcal [Eq. (43), $T = 1$]; cheat, Wiener filter with known calibration; MCMC, Markov chain Monte Carlo sampling. The gray shaded region represents the 1σ uncertainty of the CURE method.

under the usage of the resulting calibration estimate and its uncertainty afterwards. Therefore, the gradient and Hessian of the Hamiltonian $\mathcal{H}(d,\gamma) = -\ln \int \mathcal{D}s\mathcal{P}(d,\gamma,s)$ have to be calculated to find the MAP estimate $\gamma^\star$ and its uncertainty Δ, given by

$$\left.\frac{\partial \mathcal{H}(d,\gamma)}{\partial \gamma_a}\right|_{\gamma=\gamma^\star} = 0 \quad \text{and} \quad \left.\frac{\partial^2 \mathcal{H}(d,\gamma)}{\partial \gamma_a \partial \gamma_b}\right|_{\gamma=\gamma^\star} \equiv \Delta^{-1}_{ab}. \quad (38)$$

By following Ref. [1], but skipping here the full derivation, we obtain the resulting calibration formula,

$$\begin{aligned} \gamma^\star &= \Delta h, \\ \Delta^{-1}_{ab} &= \Gamma^{-1}_{ab} + \mathrm{tr}[(mm^\dagger + TD)M^{ab}] \quad \text{and} \\ h_b &= m^\dagger j^b - \mathrm{tr}[(mm^\dagger + TD)M^{ab}] \quad \text{with} \\ T &= \begin{cases} 1 & \text{for new selfcal,} \\ 0 & \text{for classic selfcal.} \end{cases} \end{aligned} \quad (39)$$

Note that the Wiener filter signal estimate $m = m(\gamma^\star)$ and its uncertainty $D = D(\gamma^\star)$ still depend on the calibration and thus Δ of Eq. (39) is not exactly the one of Eq. (38) [1]. For further details, as well as an extensive discussion of the selfcal methods, we want to point to Ref. [1].

IV. NUMERICAL EXAMPLE

A. Setup and results

To demonstrate the efficiency of the derived CURE approach we address the illustrative, simplistic, one-dimensional example used in Ref. [1] and perform a direct comparison to the selfcal schemes and MCMC sampling (see, e.g., Ref. [8]). There, a measurement device with a perfect pointlike response scans a signal field s over a (periodic) domain $\Omega = \{x\}_x = [0,1) \subset \mathbb{R}$ within a time $t \in [0,1) \subset \mathbb{R}$, but with a time-dependent calibration uncertainty, given by the calibration coefficients γ_t. This instrument exhibits a sampling rate[7] of $1/\tau = 80$ so that the ith data point, measured at time $t = i\tau$, is related to the signal at position $x_t = i\tau$. During the measurement process spatial and temporal coordinates are aligned and the data are given by

$$d_t = R_{tx} s_x + n_t = (1+\gamma_t)\delta(x - x_t)s_x + n_t, \tag{40}$$

where the signal, measurement noise n_t, and calibration coefficients γ_t are Gaussian with $\mathcal{G}(s,S)$, $\mathcal{G}(n,N)$, and $\mathcal{G}(\gamma,\Gamma)$ the corresponding PDFs with related covariance matrices S, N, and Γ. These are assumed to be known[8] and might be described by their respective power spectra in Fourier space. Following Ref. [1] we use

$$\begin{aligned} \mathcal{P}_s(k) &= \frac{a_s}{[1+(k/k_s)^2]^2}, \\ \mathcal{P}_\gamma(w) &= \frac{a_\gamma}{[1+(w/w_\gamma)^2]^2}, \quad \text{and} \\ \mathcal{P}_n(w) &= a_n. \end{aligned} \tag{41}$$

By Eq. (41) the amplitudes $a_s = \sigma_s^2\lambda_s$, $a_\gamma = \sigma_\gamma^2\tau_\gamma$, and $a_n = \sigma_n^2\tau_n$ with related variances $\sigma^2_{s,\gamma,n}$ and correlation lengths $\lambda_s = 4/k_s$, $\tau_\gamma = 4/\omega_\gamma$, and $\tau_n = \tau$ have been introduced. Within the numerical implementation we use the values $\sigma_s = 1, \sigma_\gamma = 0.3$, $\sigma_n = 0.5$, $\lambda_s = 0.3$, and $\tau_\gamma = 1.5$. This means we get a unit variance signal with calibration uncertainty of 30% and noise of 50%, which is still white (percentage values with respect to the typical signal strength).

Relating to Ref. [1], we also introduce so-called absolute calibration measurements to have additional information about the calibration that is beneficial to break the global degeneracy of the data with respect to signal and calibration variations. This means, we switch off the signal for four particular times $t_i \in \{0, 0.25, 0.5, 0.75\}$, where the calibration has the strength $c = 4$. Here, the data point d'_{t_i} is given by

$$d'_{t_i} = (1+\gamma_{t_i})c + n'_{t_i}. \tag{42}$$

During these measurements we assume the same noise statistics as before, $n' \hookleftarrow \mathcal{G}(n', N)$.

Including the absolute calibration measurements, the iterative selfcal equations [Eq. (39)] become [1]

$$\begin{aligned} \gamma^\star &= \Delta h, \\ \Delta^{-1}_{tt'} &= \Gamma^{-1}_{tt'} + \sigma_n^{-2}\delta_{tt'}\left(q_t + c^2\sum_i \delta_{tt_i}\right), \\ h_t &= \sigma_n^{-2}\left(d_t m_{x_t} - q_t + c^2\sum_i \delta_{tt_i} d'_i\right), \quad \text{and} \\ q_t &= m^2_{x_t} + T\, D_{x_t x_t} \quad \text{with} \\ T &= \begin{cases} 1 & \text{for new selfcal,} \\ 0 & \text{for classic selfcal.} \end{cases} \end{aligned} \tag{43}$$

To apply the CURE approach including the absolute calibration measurements we have to solve the ordinary differential equation of first order, according to Eq. (33) or Eq. (36), depending on whether the *zero-point* or *reference field* expansion is used. We present here the more general, but more complex, version of the reference field expansion, Eqs. (35) and (B2),

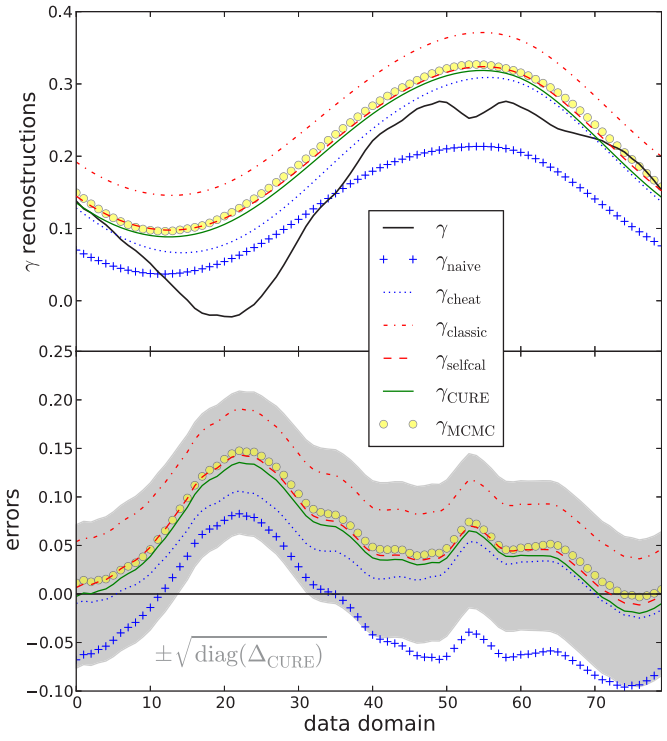


FIG. 4. (Color online) Calibration reconstructions and related errors of different approaches using Eq. (43). The terminology used follows Fig. 3. The gray shaded region represents the 1σ uncertainty of the CURE method. The reconstruction of the cheat method is not perfect, because Eq. (43) uses the Wiener filtered data (assuming the correct calibration, but non-negligible noise). The relatively good result of the naive method is a pure coincidence.

[7] Since this work is supposed to be a proof of concept we work with explicit matrices and tensors, whereby we have to limit the size of the problem for computational reasons. Further investigations are needed on how to transform this into a method using implicit tensors, and therefore suitable for "big data" problems.

[8] In case they are unknown there exist well-known methods which are able to extract the correlation structure simultaneously from data; see, e.g., Ref. [4].

because this version is constructed to deal with a larger uncertainty of the calibration than the zero-point expansion. To solve Eq. (36) we use the ordinary differential equation solver of SCIPY [9] with integrator settings: VODE, method = ADAMS. All numerical calculations have been performed using NIFTY [10,11].

Figures 3 and 4 show typical results for signal and calibration reconstruction, respectively.

Figure 5 and Tables I, II, and III show the squared error averages of the different calibration methods according to Eq. (44) at a given number of realizations[9] for signal and

[9]Note that for the statistics of 500 realizations we use a four times coarser sampling rate.

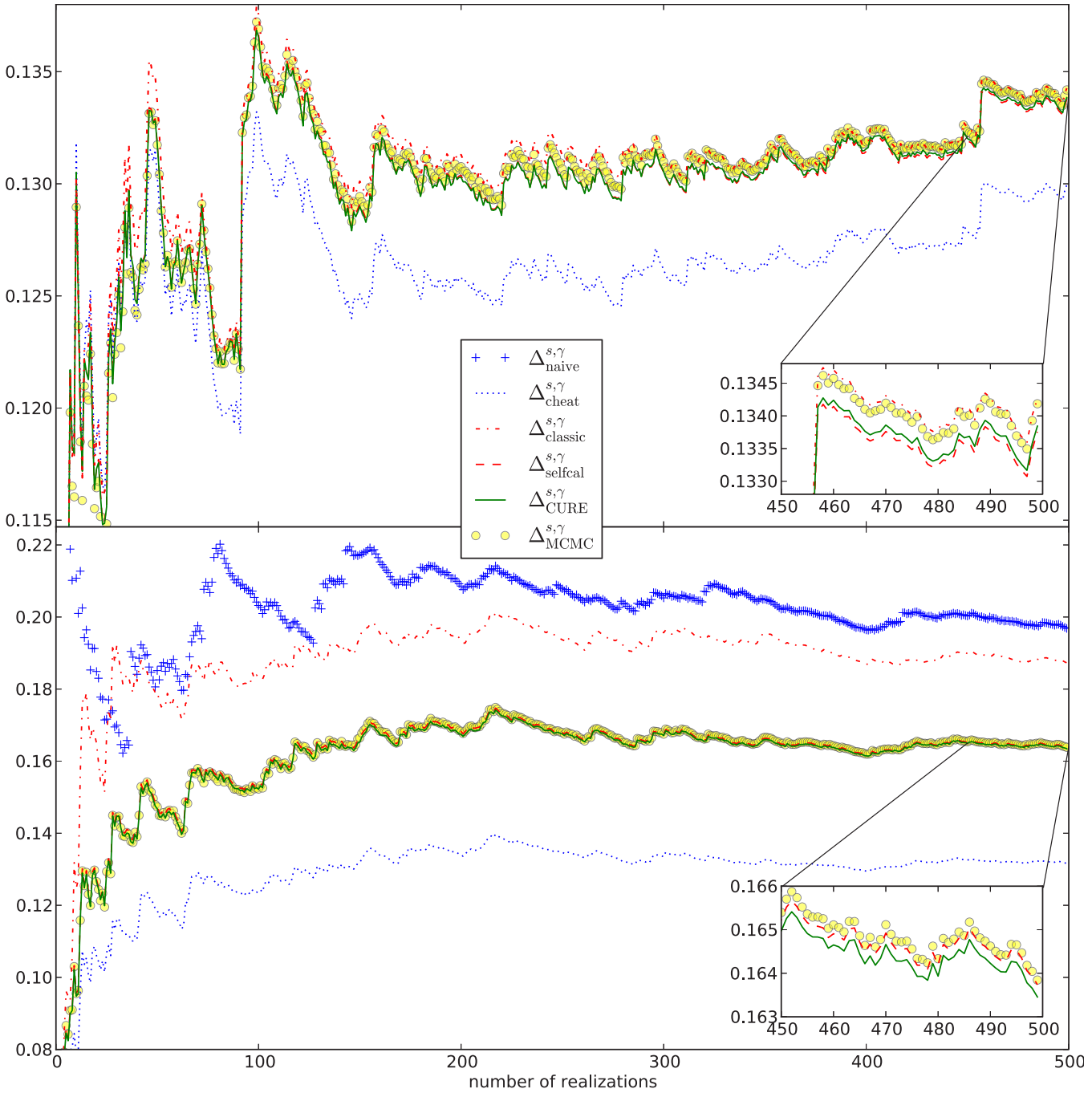


FIG. 5. (Color online) Squared error averages according to Eq. (44) at a given number of realizations for signal (upper panel) and calibration (lower panel). The best and worst results for signal and calibration yield the cheat and naive methods, respectively. In the signal domain (upper panel) all three advanced methods are very close to each other, although there is a slight preference for the CURE and selfcal methods followed by MCMC and classic methods. The results of the naive method are beyond the range of the upper panel. For the inference of calibration (lower panel) the CURE and selfcal methods perform clearly better than the classic and very similarly to MCMC calculations (see Tables I, II, and III).

TABLE I. Squared errors of signal and calibration for all methods, averaged over 500 realizations; see Fig. 5.

i	Δ_i^s	Δ_i^γ
naive	0.1635	0.1968
cheat	0.1300	0.1316
classic	0.1343	0.1873
selfcal	0.1338	0.1637
CURE	0.1338	0.1635
MCMC	0.1342	0.1638

calibration, where the following terminology is used:

$$\begin{aligned}\Delta_i^s &\equiv\langle(s-m_i)^\dagger(s-m_i)\rangle_{(d,s,\gamma)},\\ \Delta_i^\gamma &\equiv\langle(\gamma-\gamma_i)^\dagger(\gamma-\gamma_i)\rangle_{(d,s,\gamma)},\end{aligned} \tag{44}$$

where i indicates naive, cheat, classic, selfcal, CURE, or MCMC, referring to the Wiener filter methods without and with information about calibration, the classical and new selfcal schemes, the CURE scheme, and MCMC sampling, respectively.

B. Discussion

As Figs. 3, 4, and in particular Fig. 5 with the related Tables I, II, and III illustrate, the CURE and new selfcal (selfcal) approach prevail against classical selfcal (classic) and Wiener filtering with unknown calibration (naive) and perform similarly to the MCMC method. The last represents in principle the best method by avoiding any approximations, but it is also the most expensive one. Its small underperformance in comparison to CURE and selfcal has its origin in using still insufficient samples for the MCMC chains[10] to converge fully. Increasing their number, however, would increase the numerical effort significantly.

The upside of CURE is that it is not iterative since it involves only the solution of a single system of coupled ordinary differential equations (ODEs). For ODEs, in turn, there exist a number of numerical solvers with adaptive step-size control that work well, which might save significant amounts of computational time.[11] This is, however, only true if one finds a clever implementation or sparse representation of $\Lambda^{(3)}$, because the term $\Lambda^{(3)}_{yzv}D_{zv}$ required in Eq. (35) might become a bottleneck within a calculation due to its complex correlation structure (in contrast to the $\Lambda^{(4)}$ term). Another downside is the higher level of complexity in comparison to the new selfcal that naturally arises with a renormalization calculation.

[10]For each signal realization we have to run a separate chain. In the numerical example used in this work, a single chain consists of 2×10^3 independent samples.

TABLE II. Improvements of the methods' signal squared errors with respect to the naive method, averaged over 500 realizations.

i	$\Delta^s_{\text{naive}}-\Delta^s_i$	Improvement
naive	0.0000	0.00%
cheat	0.0335	100.00%
classic	0.0292	87.16%
selfcal	0.0297	88.86%
CURE	0.0297	88.66%
MCMC	0.0293	87.46%

TABLE III. Improvements of the methods' calibration squared errors with respect to the naive method, averaged over 500 realizations.

i	$\Delta^\gamma_{\text{naive}}-\Delta^\gamma_i$	Improvement
naive	0.0000	0.00%
cheat	0.0652	100.00%
classic	0.0095	14.57%
selfcal	0.0331	50.77%
CURE	0.0333	51.07%
MCMC	0.0330	50.61%

V. CONCLUDING REMARKS

We derived a calibration-uncertainty renormalized estimator method to infer a signal and consequently a calibration without knowledge of the calibration but of its covariance. The basic idea of the CURE method is to perform a perturbation calculation around a reference field, an *a priori* determined reconstruction of the signal without knowledge of the calibration. "Perturbatively" means that we successively take into account higher-order terms of calibration uncertainty. In this way, the problem of signal reconstruction without knowledge of the calibration, which is often solved by iterative or brute-force sampling methods, becomes a single system of ordinary differential equations.

We applied the method to a mock example and compared it against other existent calibration methods. For this example we found that the CURE method performs extremely similarly to the new selfcal and MCMC sampling, and clearly beats the Wiener filter without calibration as well as the classical selfcal method in terms of reconstruction accuracy. Although it obviously performs well, a recommendation to favor this method over the new and classical selfcal methods depends on the particular problem at hand as well as on the numerical implementation, as discussed in Sec. IV B. Therefore it serves as an alternative to them.

ACKNOWLEDGMENTS

Calculations were realized using the NIFTY [10,11] package as well as SCIPY [9].

[11]We want to mention that we also found realizations for certain levels of signal, noise, and calibration where the SCIPY solver did not converge. The reason for this might be an initial guess too far away from the correct solution. In many cases one could cope with this problem by significantly reducing the step size δt or using the optimization described in Sec. III A 3. We did not elaborate on this, however, since this work is supposed to be a proof of concepts.

APPENDIX A: FEYNMAN RULES

The Feynman rules originally stated in and inherited from Ref. [4] read as follows:

(1) Open ends of lines in diagrams correspond to external coordinates and are labeled by such. Since the partition sum in particular does not depend on any external coordinate, it is calculated only from summing up closed diagrams. However, the field expectation value $m(x) = \langle s(x)\rangle_{(s|d)} = \delta \ln[\mathcal{Z}(d,J)]/\delta J(x)|_{J=0}$ and higher-order correlation functions depend on the coordinates and therefore are calculated from diagrams with one or more open ends, respectively.

(2) A line with coordinates x' and y' at its ends represents the propagator $D_{x'y'}$ connecting these locations.

(3) Vertices with one leg get an individual internal, integrated coordinate x' and represent the term $j_{x'} + J_{x'} - \Lambda^{(1)}_{x'}$.

(4) Vertices with n legs represent the term $-\Lambda^{(n)}_{x'_1\ldots x'_n}$, where each individual leg is labeled by one of the internal coordinates $x'_1,\ldots,x'_n$. This more complex vertex structure, as compared to quantum field theory, is a consequence of nonlocality in IFT.

(5) All internal (and therefore repeatedly occurring) coordinates are integrated over, whereas external coordinates are not.

(6) Every diagram is divided by its symmetry factor, the number of permutations of vertex legs leaving the topology invariant, as described in any book on field theory.

APPENDIX B: RENORMALIZATION FLOW EQUATIONS INCLUDING ABSOLUTE CALIBRATION MEASUREMENTS

This section derives the generalization of the renormalization flow equations in the presence of absolute calibration measurements. These measurements can be included in the prior knowledge of the calibration coefficients, $\mathcal{P}(\gamma) = \mathcal{G}(\gamma - m_\gamma, D^\gamma)$, with $(m_\gamma)_a$ the Wiener filter solution for γ with uncertainty D^γ using the absolute calibration measurements only. Hence, the likelihood becomes

$$\mathcal{P}(d|s) = \int \mathcal{D}\gamma\, \mathcal{P}(d|s,\gamma)\mathcal{G}(\gamma - m_\gamma, D^\gamma) = \mathcal{G}\left(d - \check{R}s, N + \sum_{ab} D^\gamma_{ab} R^a s s^\dagger R^{b\dagger}\right),$$
$$\check{R} \equiv R^0 + \sum_a (m_\gamma)_a R^a. \tag{B1}$$

Compared to the result without measurements of absolute calibration, Eq. (23), the response R and the calibration covariance have been replaced by $\check{R}$ and D^γ, respectively. This means that the response became modified by new, additional, information from the absolute calibration measurements and associated with the uncertainty D^γ, which is not diagonal anymore. The resulting reference field expansion of the Hamiltonian, Eq. (34), yields the following assignments:

$$\check{D} = (S^{-1} + \check{R}^\dagger N^{-1}\check{R})^{-1},\ \check{j} = \check{R}^\dagger N^{-1} d,\ \check{m} = \check{D}\check{j},\ M^{\check{x}} \equiv \check{R}^\dagger N^{-1} R^x,$$
$$\Lambda^{(1)}\phi = \frac{1}{\delta t}\underbrace{(\check{m}^\dagger \check{D}^{-1} - \check{j}^\dagger)}_{=0}\phi + \sum_{ab} D^\gamma_{ab}\Big\{\check{m}^\dagger M^{ab}\phi - \frac{1}{2}j^{a\dagger}(\phi\check{m}^\dagger + \check{m}\phi^\dagger)j^b - \frac{1}{2}\check{m}^\dagger M^{\check{a}}(\phi\check{m}^\dagger + \check{m}\phi^\dagger)M^{\check{b}}\check{m}$$
$$- \frac{1}{2}\check{m}^\dagger M^{\check{a}}\check{m}\check{m}^\dagger M^{\check{b}}\phi + j^{a\dagger}\check{m}\check{m}^\dagger M^{\check{b}}\phi + \frac{1}{2}j^{a\dagger}(\phi\check{m}^\dagger + \check{m}\phi^\dagger)M^{\check{b}}\check{m} + \frac{1}{2}\check{m}^\dagger M^{\check{a}}(\phi\check{m}^\dagger + \check{m}\phi^\dagger)j^b\Big\},$$
$$\Lambda^{(2)}[\phi,\phi] = \frac{1}{2}\sum_{ab} D^\gamma_{ab}\{\phi^\dagger M^{ab}\phi - j^{a\dagger}\phi\phi^\dagger j^b - \phi^\dagger M^{\check{a}}\check{m}\check{m}^\dagger M^{\check{b}}\phi - \check{m}^\dagger M^{\check{a}}\phi\phi^\dagger M^{\check{b}}\check{m} - \phi^\dagger M^{\check{a}}(\phi\check{m}^\dagger + \check{m}\phi^\dagger)M^{\check{b}}\check{m}$$
$$- \check{m}^\dagger M^{\check{a}}(\phi\check{m}^\dagger + \check{m}\phi^\dagger)M^{\check{b}}\phi + j^{a\dagger}(\phi\check{m}^\dagger + \check{m}\phi^\dagger)M^{\check{b}}\phi + \phi^\dagger M^{\check{a}}(\phi\check{m}^\dagger + \check{m}\phi^\dagger)j^b$$
$$+ j^{a\dagger}\phi\phi^\dagger M^{\check{b}}\check{m} + \check{m}^\dagger M^{\check{a}}\phi\phi^\dagger j^b\} + 1\text{ perm},$$
$$\Lambda^{(3)}[\phi,\phi,\phi] = -\sum_{ab} D^\gamma_{ab}\Big\{\frac{1}{2}\phi^\dagger M^{\check{a}}(\phi\check{m}^\dagger + \check{m}\phi^\dagger)M^{\check{b}}\phi + \frac{1}{2}\check{m}^\dagger M^{\check{a}}\phi\phi^\dagger M^{\check{b}}\phi + \frac{1}{2}\phi^\dagger M^{\check{a}}\phi\phi^\dagger M^{\check{b}}\check{m}$$
$$- \frac{1}{2}j^{a\dagger}\phi\phi^\dagger M^{\check{b}}\phi - \frac{1}{2}\phi^\dagger M^{\check{a}}\phi\phi^\dagger j^b\Big\} + 5\text{ perm},$$
$$\Lambda^{(4)}[\phi,\phi,\phi,\phi] = -\frac{1}{2}\sum_{ab} D^\gamma_{ab}\phi^\dagger M^{\check{a}}\phi\phi^\dagger M^{\check{b}}\phi + 23\text{ perm}. \tag{B2}$$

[1] T. A. Enßlin, H. Junklewitz, L. Winderling, M. Greiner, and M. Selig, Phys. Rev. E **90**, 043301 (2014).

[2] T. A. Enßlin, M. Frommert, and F. S. Kitaura, Phys. Rev. D **80**, 105005 (2009).

[3] C. Osborne, In. Stat. Rev./Rev. Int. Stat. **59**, 309 (1991).

[4] T. A. Enßlin and M. Frommert, Phys. Rev. D **83**, 105014 (2011).

[5] T. Bayes, Philos. Trans. R. Soc. London **53**, 370 (1763).

[6] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series* (Wiley, New York, 1949).

[7] S. L. Bridle, R. Crittenden, A. Melchiorri, M. P. Hobson, R. Kneissl, and A. N. Lasenby, Mon. Not. R. Astron. Soc. **335**, 1193 (2002).

[8] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[9] Computer code SCIPY, http://docs.scipy.org/doc/scipy/reference/generated/scipy.integrate.ode.html.

[10] Computer code NIFTY, http://www.mpa-garching.mpg.de/ift/nifty.

[11] M. Selig, M. R. Bell, H. Junklewitz, N. Oppermann, M. Reinecke, M. Greiner, C. Pachajoa, and T. A. Enßlin, Astron. Astrophys. **554**, A26 (2013).