I. INTRODUCTION

Modeling the dynamic evolution of systems has always been at the very core of many scientific areas. While fundamental classical evolution is fully determined via ordinary and partial differential equations, many real-world models only adapt an effective probabilistic approach. Stochastic differential equations (SDE) provide an excellent way to construct such models for many dynamical systems. Generally SDEs consist of two constituents, a part describing the deterministic evolution of the system and another part describing stochastic evolution. The presence of a stochastic term in the equation can have a variety of reasons. It may be that the prescribed system is influenced by an external quantity which can only be described up to its statistical properties. Or the fundamental deterministic evolution is too complex to be fully modeled, for example due to non-linear self interactions or unknown inhomogeneous couplings, and should be approximated via a simpler (e.g. linear and homogeneous) deterministic evolution. All other influences can be approximated in terms of a stochastic constituent.

As a consequence, SDEs provide a great flexibility in order to decide up to which level of detail an effective description should capture deterministic evolution, and at which level a probabilistic description is sufficient. In the past, SDEs have been extensively used to simplify or approximate dynamic evolution, e.g. when modeling Brownian motion [1], in the Langevin dynamics [2] or the Black Scholes model [3]. On the other hand SDEs can also be used in order to extract information about the dynamic evolution of a system from observations. The possibility to add stochastic constituents to SDEs is useful to decide a priori up to which level dynamic evolution is modeled. However, inference of SDEs is complicated due to a degeneracy between the two constituents, the deterministic and the stochastic term. As an illustrative example, consider a case of a simple random walk in time with an additional unknown linear drift. From observations we know which values the system obtained at several moments in time. Given only this incomplete information about a single realization of the stochastic process the reconstruction problem would be completely degenerate as it is impossible to decide whether observed temporal changes are due to the drift or due to the random walk. Only due to additional prior information such as stationarity of the drift it is possible to arrive at a non-degenerate interpretation. Degeneracies can appear on all levels of the inference and therefore prior modeling of the considered problem at hand is inevitable. In many applications this problem is solved by restricting a priori possible dynamics via parametrization with only a few free parameters which have to be inferred [4]. However under certain conditions it is also possible to perform non-parametric inference. In [5] it was shown that for a linear autonomous system the spectrum of the corresponding Green’s function, or propagator, can be constructed non-parametrically given a mild regularization on the spectrum as well as the statistics of the stochastic constituent. We will call these constituents excitations in the following.

The spectrum of the Green’s function determines how the field reacts to excitations. The delay of the reaction of a system to an excitation is however still ambiguous. This potential time lag can be encoded in the phase of the Green’s function, which therefore needs to be inferred as well. This appears to be problematic since in many realistic settings the statistical properties of the excitations can only be constrained up to the expected statistical amplitude, but it is not possible to a priori constrain the phase configuration of excitations. In this work we show which conditions still allow to reconstruct the full dynamical process in one-dimensional (only temporal) as well as multi-dimensional (spatio-temporal) settings. In particular we show that two fundamental concepts, namely locality and causality, appear to be sufficient to construct a prior for the dynamic Green’s function that ultimately allows for a non-degenerate inference of it.

This paper structures as following: In section [1] we
outline some basic properties of linear autonomous SDEs which are needed for the remaining discussion and section III introduces the prior assumptions that enter the inference problem. In section IV we outline a resulting inference algorithm given a generic measurement problem which is applied to several mock data examples in section V. Ultimately we conclude the paper in section VI with a brief summary of the results and an outlook to possible applications and extensions.

II. FIELD AND DYNAMICS

We are interested in reconstructing a field \( s \), defined over space and time, and its dynamical law, governing its evolution in space-time. We assume that \( s \) solves a linear stochastic differential equation (SDE) of the form

\[
\mathcal{L} s = \xi \, ,
\]

where \( \mathcal{L} \) is a linear differential operator, and \( \xi \) denotes an unknown excitation field with assumed to be known statistics \( P(\xi) \). If we restrict the analysis to homogeneous and autonomous processes then \( \mathcal{L} \) is a homogeneous operator, i.e.

\[
\mathcal{L}_{xx'} = q(\partial_{x'}) \delta_{xx'} \, .
\]

Here \( x = (t, x) \) is a space-time vector, \( q(\partial_{x'}) \) is a differential operator encoding function, and \( \mu \) is an index of the coordinates.

Fourier transforming eq. [1] by applying \( F \) with

\[
F_{xk} = e^{ikx}
\]

to it, and using eq. [2] yields the Fourier space representation of the differential equation

\[
q(ik)s_k = \xi_k \, ,
\]

which can be solved trivially

\[
s_k = \frac{\xi_k}{f_k} \, ,
\]

where we also introduced the complex valued differential operator encoding field \( f_k := q(ik) \). Note that if we want the model (eq. [1]) to be real, both fields \( \xi_k \) and \( f_k \) have to be hermitian.

Our goal is to infer \( s \) and its governing differential equation from observations of \( s \) alone. These observations may be noisy, incomplete and / or are retrieved via a complicated (possibly non-linear) instrument response. Consequently we rely on Bayesian inference in order to reconstruct the desired quantities. We note that the differential equation is fully characterized by \( f \) and \( \xi \) and we can also calculate the corresponding solution from them using eq. [3]. This indicates that we can reformulate the problem such that we only aim to reconstruct \( f \) and \( \xi \) from observations.

Before we can solve the corresponding measurement equation, we have to construct a meaningful prior for \( f \). Up to this point, the statistics of \( \xi \) remain known.

III. DIFFERENTIAL OPERATOR ENCODING FIELD PRIORs

Physical processes that follow an autonomous differential equation usually are described in terms of a low order polynomial in the derivatives. While we do not want to restrict the analysis to such structures, we note that this property serves as an ingredient for the prior of \( f \). In particular we aim to impose a certain degree of smoothness for \( f \). What we can do is to construct a prior for \( f \) such that it punishes second (or any other) order derivatives of the field. One way to do so is to propose a Gaussian prior

\[
P(f) = G(f, T) = \frac{1}{|2\pi T|^2} e^{-\frac{1}{2} f^T T^{-1} f} \, ,
\]

where the exponential factor reads

\[
-\frac{1}{2} f^T T^{-1} f = -\frac{1}{2\sigma^2} \int |\Delta_k f_k|^2 dk \, .
\]

For details see [6].

A. Symmetries

Due to the fact that we aim to reconstruct \( \xi \) and \( f \) from observations on \( s \) alone we notice that a corresponding likelihood will be invariant under simultaneous manipulations of these quantities. I.e.

\[
s_k = \frac{\xi_k}{f_k} = \frac{\xi_k}{f_k} \frac{r_k}{r_k} = \frac{\xi_k}{f_k} a_k e^{i\phi_k} = \frac{\xi_k}{f_k} e^{i\phi_k} = \frac{\xi_k}{f_k} \, ,
\]

where we split the manipulations \( r_k \) into its amplitude \( a_k \) and its phase \( \phi_k \) as we will treat them separately.

If we aim to use the effective model (eq. [4]) only to construct a prior for the observed quantity \( s \), there is no need to break this symmetry since by definition, it leaves the resulting \( s \) invariant. In this case the relation between \( \xi \) and \( s \) is only artificial, it is only used to construct a prior for \( s \) and we do not care if such an interaction is actually realizable in reality.

This picture completely changes once we aim to find a physical interpretation of \( f \) and \( \xi \) from observations of \( s \). Although implemented as an effective model, there exist possible physical interpretations of the excitation field as an external force applied to the system, or a non-linear self-interaction of the system. Therefore, it is of great interest whether or not it is possible to \textit{a posteriori} analyze the resulting reconstruction regarding these questions. This is obviously not possible if there is a degeneracy between \( f \) and \( \xi \) as the reconstructed configuration is not unique. Therefore, in applications where the effective model we aim to infer should permit a specific physical interpretation, we need to break this degeneracy via priors which are not invariant under such transformations. Fortunately, the physical interpretation of \( \xi \) and \( f_k \) provides the necessary prior constraints.
B. Amplitude degeneracy

In many applications the expected amplitude of the excitations can be assumed to be known (i.e. if the spectrum of $\xi$ is known). In addition, we require $f_k$ to be a smooth function of $k$ which already reduces the amplitude degeneracy. Throughout this work we restrict ourselves to such cases.

C. Phase degeneracy

The degeneracy resulting from phase manipulations is somewhat more problematic, since in many applications we cannot set the phases of the excitations. We notice that if the prior of $\xi$ is assumed to be Gaussian distributed with a homogeneous covariance, such phase rotations wont affect it in any way. I.e. 

$$P(\xi') = P(\xi)$$

$$\xi_k' = \xi_k e^{i\phi_k}.$$  

In order to resolve this we need to provide additional information to the reconstruction problem. One trivial way would be to provide direct, additional measurements of the excitation, which we can use to fix the phase of $\xi$. However, in many realistic settings such measurements are difficult or even impossible to receive.

In cases where the statistics of $\xi$ can be characterized via a distribution which is not invariant under local phase transformations, this prior knowledge also breaks the invariance of the posterior distribution, allowing for a phase sensitive reconstruction. For example if the excitations are assumed to be sparse, or if we deal with inhomogeneous distributed excitations.

Depending on the strength of such priors, or the quality of possible direct measurements, these approaches (partially) resolve the phase degeneracy. However, there exist also many applications in which we have to rely on less informative priors of the excitations, e.g. if we want to model multiple unknown external interactions or even unknown non-linear self-interactions via excitations.

Consequently, if we cannot break the phase degeneracy via manipulating $\xi$, we seek to find a way to constrain $f$ such that the phase configuration is unique. To some extend this is already achieved via the smoothness prior. As we can see in Figure[2] the locality of a response of the system is related to the order in the derivatives of the operator. Since lower order derivatives are a priori favoured, the most local configuration that is consistent with the data is the one selected during reconstruction. This is perfectly motivated physically, since we know that for direct (classical) interactions, locality is a fundamental concept. Consequently apparent non-localities (realized via higher order derivatives) indicate additional hidden layers within the system, since the excitation does not influence the observed quantity $s$ directly, but couples to an additional unobserved quantity which then couples to $s$. Such a configuration is of course possible and perfectly realizable, but if the system can also be described via a more direct interaction, the corresponding model should be favoured. This is achieved via our prior.

Although the smoothness prior determines localization, it does not set the temporal direction of propagation, as can be seen in Figure[2]. Therefore the phase degeneracy corresponding to the direction of propagation remains and needs to be fixed. This is a well known degeneracy when constructing a response (Greens function) of a physical system. In general there exist both solutions, the advanced and the retarded Greens function as a response of a system, which could appear as a superposition in the reconstructed response. This issue can be addressed once we add causality as another physical principle besides locality. While it remains a topic to be discussed whether the excitations caused a response of the system or vice versa, a superposition of causal (retarded) and anti-causal (advanced) responses usually does not make any sense.

In order to fix this degeneracy we follow the standard way of obtaining retarded or advanced responses, by convolving with a step function $\Theta$ in time. This reads

$$G_{\text{ret./adv.}} = \left( F \Theta(\pm t) F^{-1} \frac{1}{f} \right)_k,$$  

where $\Theta(\pm t)$ denotes a diagonal operator in space and time with $\Theta(\pm t)$ on its diagonal.

In principle we are free to choose if $\xi$ should be the cause or the effect, but we notice that the way we constructed our model clearly indicates that $\xi$ should cause the system to respond and not vice versa. We also want to emphasize at this point that this is a choice we make and the direction of causality is not part of our inference. I.e. we do not aim to perform causal inference, we rather define that $\xi$ causes $s$.

Consequently we obtain a modified relation for our solution of $s$:

$$s_k = G_k(f) \xi_k,$$  

where we chose $G \equiv G_{\text{ret.}}$. Retardation trivially ensures that no anti-causal response of the system is allowed.

Ultimately we are left with one global phase degeneracy ($\phi_k \equiv \phi$). Noticing the requirement that the observed process (eq. [1]) should be real, the global phase also has to be real. We get

$$e^{i\phi} = e^{im\pi} = \pm 1, \quad n \in \mathbb{N}.$$  

In other words, the only remaining degeneracy is an overall multiplication of the r.h.s. of eq. [1] with a minus sign.

D. Relation to minimum-phase

There are some structural similarities of our prior to a well known signal processing approach to phase configurations, namely the minimum-phase filter [7].
FIG. 1. Excitation $\xi$ and responses of various systems defined in terms of powers of $f$, with $f_t = \partial_t + \gamma$ and $\gamma$ small. We see that higher order derivatives result in apparent non-local responses. Since time derivatives are the generators of temporal translation, the case where the response is the exponential of $-\Delta t f$ leads to translations by $\Delta t$.

In signal processing, one often faces the problem of constructing a meaningful LTI (linear translation invariant) filter response, given a corresponding spectrum. A minimum phase filter can be constructed by relating the cepstrum (Fourier transformed logarithmic spectrum), with the phase via

$$\phi_\omega = F_{\omega t} \text{sign}(t) C_t,$$

where

$$C_t = -F_{t^{-1}} \log(|f_\omega|)$$

(15)
denotes the cepstrum. A minimum phase filter is defined to be the causal and stable LTI filter with the least amount of group delay, corresponding to a given spectrum. This indicates that the resulting filter response is maximally localized in time. Here we notice the similarities to our prior: The smoothness prior favors more local and stable responses and retardation ensures causality of the response. However, in our approach these restrictions are formulated as a priori assumptions and therefore the posterior solution can deviate from these assumptions if there is enough evidence in the data to support this. This is not possible in a minimum-phase setup as the phase configuration is uniquely defined via the spectrum. In addition the inverse of the response of a minimum phase configuration is also causal and stable. While this is a desirable property in control systems, where the minimum phase filter is commonly used, we do not aim to make a statement on stability of the inverse of the response since there is not necessarily a meaningful physical interpretation of this response. This inverse should not be confused with a back-reaction of the system on excitations. These should also be causal and stable but have to be distinguished since they might have their own Green’s function.

It is possible to incorporate the minimum phase condition as a forward model to construct corresponding propagators. This yields

$$G^{MP} = \exp \left[ - \left( 1 + F \text{sign}(t) F^{-1} \right) \log(|f|) \right],$$

(16)

with $f$ being a priori distributed according to the smoothness prior given by eq. 6.

E. Light cone prior

So far, we restricted the discussion of causality to the temporal domain. As we generally aim to reconstruct the space-time evolution of a field, we notice that the physical concept of causality also implies a maximal finite propagation speed of interactions. This leads to propagation within a light cone, as depicted in figure 2. This information can be used to further restrict the reconstruction to propagators that are within the light cone, given a maximal propagation speed $c$. We can implement this constraint by rewriting eq. 11 as

$$G = F \Theta(t) \Theta(t^2) F^{-1} \frac{1}{f},$$

(17)

where

$$t^2 = t^2 - x^\dagger C^{-1} x.$$
This ensures that the propagator is only non-zero within the light cone. In general we might expect different maximal propagation speeds in different directions, and consequently $C$ becomes a symmetric tensor. In case propagation is isotropic in space with speed $c$ we have $C_{ij} = c^2 \delta_{ij}$. Note that even in cases where we do not know $C$, for example if the medium in which the interaction is realized is unknown, such a constraint can also be useful if we elevate $C$ to be an additional unknown parameter that has to be inferred. Since $C$ is the same for all scales, an inference algorithm can use large scale information, where the signal to noise ration (SNR) is usually higher, to determine $C$, which then effectively increases the SNR for smaller scales. A detailed description how $C$ enters the reconstruction can be found in appendix A.

IV. INFEERENCE

In general we aim to reconstruct $\xi$ and the dynamical operator encoding field $f$ from observations of $s$ alone. Throughout this work we assume a generic measurement equation of the form

$$d = R[s] + n ,$$

(19)

where $d$ denotes the measured data, $R$ is an (in general non-linear) instrument response and $n$ denotes the measurement noise. Furthermore we assume signal independent Gaussian distributed noise,

$$P(n) = \mathcal{G}(n, N) ,$$

(20)

with $N$ being the known covariance structure of $n$. 

FIG. 2. Left: Causal response of a system defined as a combination of two damped harmonic oscillators with different masses. Right: Anti-causal response of the system where one of the oscillations travels backwards in time.
A. Field inference only

Let us assume that the dynamics encoding field $f$ is known to be $f$. Our goal is now to infer $\xi$ given $d$ and $f$. Note that $s$ is also inferred, it can be calculated from $f$ and $\xi$ using eq. 9. The posterior reads

$$P(\xi|d,f) \propto P(d|\xi,f)P(\xi) = \mathcal{G} \left( d - R \left[ \mathcal{F} \xi \right], N \right) P(\xi) .$$

(21)

If we assume $R$ to be a linear operator, and also assume that $\xi$ is a priori Gaussian distributed with a known covariance structure $\Xi$, the posterior distribution is also a Gaussian

$$P(\xi|d) = \mathcal{G}(\xi - m, D) ,$$

(22)

where

$$D = \left( (\mathcal{F}^{-1})^\dagger \mathcal{F}^\dagger R^\dagger N^{-1} R^T \mathcal{F}^{-1} + \Xi^{-1} \right)^{-1} .$$

(23)

Introducing $R' := R^T \mathcal{F}^{-1}$ one can write

$$m = D j = D R'^\dagger N^{-1} d$$

(24)

and

$$D = (R'^\dagger N^{-1} R' + \Xi^{-1})^{-1} .$$

(25)

This resembles the solution of the famous Wiener filter equations [8], applied to a field theoretical setting [9, 10].

B. Inference of a field and its dynamics

Now we want to infer both $\xi$ and $f$ from the data. The posterior reads

$$P(\xi, f|d) \propto P(d|\xi, f)P(\xi)P(f)$$

$$= \mathcal{G}(d - R \mathcal{F} s(\xi, f)], N) P(\xi) P(f) ,$$

(26)

where $s(\xi, f)$ is defined according to eq. [12]. Note that even for a linear measurement response $R$, the inference problem becomes intrinsically non-linear, due to the non-linear dependency of $s$ on $f$. We rely on a variational approach to solve the inference problem, where we approximate the posterior with a Gaussian distribution by minimizing the corresponding forward Kullback-Leibler divergence (KL). Here we follow the approach introduced via [11] and [12]. Introducing the vector $\theta = (\xi, f, C)$ into which we include all parameters of the inference, we approximate the posterior via a Gaussian in $\theta$ where we also reconstruct the correlation structure of $\theta$ including all cross-correlations. This reads

$$\text{KL}(\hat{P}|P) = \langle \mathcal{H}_\theta \rangle_{\hat{\theta}} - \langle \mathcal{H}_\theta \rangle_\theta ,$$

(27)

where

$$\hat{P}(\theta) = \mathcal{G}(\theta - m, D)$$

(28)

and

$$\mathcal{H}_\theta = -\ln (Q) .$$

(29)

Minimizing the KL w.r.t. the mean $m$ and the covariance $D$ leads to the desired solution. Following [11], we sample from the approximate posterior which enables us to calculate approximate posterior expectation values for any desired quantity. This is important since calculating $s$ from $\theta$ involves non-linear operations and consequently $\langle s(\theta) \rangle_\theta \neq s(m)$.

V. APPLICATION

In order to demonstrate the applicability of our method we apply it to several mock data examples. These include examples consistent with our prior assumptions as well as an inconsistent one. First, we demonstrate the performance of the algorithm in a one dimensional setting, where we only aim to infer the Green’s function of the system, given the excitations and data. In the second example we perform a reconstruction of the excitation as well as the Green’s function from data alone, in a 1+1 dimensional setting. Next, we add another level of complexity by non-Gaussian statistics in the excitations. Hereby we perform source detection, the task of inferring sparse excitations at various locations in space-time, in a case where also the Green’s function is unknown. Ultimately we apply the reconstruction algorithm to data generated from a non-linearly evolving system, to demonstrate the performance in an inconsistent setup.

All examples were implemented using the NIFTy software package [13, 14] in its version 5 [15]. This package allows to infer field quantities and to represent the remaining uncertainties via posterior samples. We included an implementation of the here introduced prior structure for the Green’s function in this publicly available package.

A. Temporal evolution

In our first application, shown in fig. 3, we aim to demonstrate the inference of the dynamics encoding field alone in a one dimensional setting where there is only a temporal evolution to be reconstructed. The excitation field is known during inference. We generate mock data according to eq. [19] with $R$ being the identity. The excitations were drawn from an inverse gamma distribution to model known, sparse excitations of the system (i.e. in a laboratory setting where the unknown system is driven via sparse excitations). The mock signal $\phi$ as well as corresponding data $d$ is shown in the top-left panel of fig. 3. The dynamical system used to generate signal and data is of the form:

$$\left( \partial_t^2 + m^2 + \gamma \partial_t \right) \left( \partial_t^2 + \tilde{m}^2 + \tilde{\gamma} \partial_t \right) \phi_t = \xi_t ,$$

(30)
FIG. 3. **Top:** On the left we depict the signal (red line), the data (blue dots) as well as 100 over-plotted posterior samples (gray). The right panel shows the mock propagator (Greens function) in the temporal domain (red) as well as corresponding posterior samples. **Bottom:** Left: Excitation field used to generate the signal. Right: Logarithmic spectrum of the mock propagator (red) and posterior samples.

with $\xi_t$ being the inverse gamma distributed excitation field and $(m, \gamma, \tilde{m}, \tilde{\gamma}) = (0.6, 0.23, 0.2, 0.03)$. The goal is now to reconstruct the differential operator encoding field $f$ given $d$ and $\xi$. The results are shown in fig. 3 as well.

We see that the reconstruction of the Green’s function behaves as expected and is in agreement with the ground truth in the temporal domain, within uncertainties. Due to the fact that the reconstructed dynamics is uncertain, the recovered signal also has uncertainty although the excitations are known. The reconstructed Green’s function indicates that it is indeed possible to reconstruct an apparent non-local response of the system (due to higher order derivatives in this setup) since the true propagator as well as its reconstruction show oscillations that grow in the beginning of the propagator before decaying exponentially. We also notice that there is relatively high uncertainty in the beginning of the propagator. This is caused by the low initial response to excitations of the true propagator. The initial part of the reconstructed propagator is purely dominated by noise and thus only constrained up to the noise level (the noise variance $\sigma_n$ is set to be $\sigma_n = 20$ and the temporal domain is discretized via 512 equidistant pixels).

We also notice that the posterior solution for the spectrum levels out for small modes (large $\omega$) below the signal to noise ratio, while the true spectrum shows a polynomial divergence (ultimately also the true spectrum levels out in the numerical example due to the finite size of the considered space). The uncertainty increases in this region, but not enough to capture the true solution. Here we notice the limitations of the variational inference used by NIFTy5 [16], which provides a local approximation of the posterior with a Gaussian. Consequently the true uncertainty might be underestimated, as in this case. However this deviation occurs orders of magnitudes below the peak of the spectrum and therefore has only a barely visible effect on the reconstruction of the signal. One way to allow for a better extrapolation to higher frequencies would be to provide a more restrictive prior for the dynamics encoding field e.g. by defining it on a polynomial basis which is a more suitable basis for this
FIG. 4. **Top:** The left panel shows the spatio-temporal masked and noisy data, drawn from the mock signal (middle panel) and the corresponding signal reconstruction (right panel). **Bottom:** Logarithmic one-sigma posterior uncertainty (left panel), logarithmic residual between the true signal and the posterior mean (middle panel), as well as an approximate posterior sample (right panel).

particular setup as the true dynamics is also described in terms of a polynomial. However we aim to provide a general and less restrictive approach here capable of also reconstructing non-polynomial dynamics encoding functions and consequently being less able to extrapolate to regions where we have no information provided by data.

**B. Space-time evolution**

In our second example, shown in fig. 4, we aim to reconstruct the dynamics as well as the excitations in a spatio-temporal (1+1 dimensional) setting from incomplete and noisy observations. We aim to infer the dynamical field \( f \), the propagation speed \( c \), and the excitations \( \xi \) from measurements \( d \) of \( \phi \) alone. The dynamical system used for the generation of mock data is a product of a damped harmonic oscillator and an advection-diffusion generating term. This reads

\[
\mathcal{L} = (\partial_t^2 - c^2 \partial_x^2 + m^2 + \alpha \partial_t - \beta \partial_x) (\partial_t + \tilde{m}^2 - \gamma \partial_x^2 + \tilde{\beta} \partial_x).
\]

The signal \( \phi \), the data \( d \) and the reconstruction of \( \phi \) are shown in fig. 4 for a case with \((c, m, \alpha, \beta, \tilde{m}, \gamma, \tilde{\beta}) = (0.4, 0.16, -0.19, -0.05, 0.1, 0.5, 0.2)\). We generate the data according to eq. 19 with a linear measurement response which partially \((\approx 25\%)\) masks the observed region, and Gaussian distributed noise with \( \sigma_n = 10 \). The space-time is discretized via a regular grid with \( 256 \times 200 \) pixels, respectively.

The reconstruction algorithm is capable of reconstructing the signal in regions where we have observations thereof, while being relatively blind in unobserved regions. Consequently the posterior uncertainty is higher there. In addition, we notice that unlike the posterior mean, posterior samples consistently fill unobserved regions. Although in these regions the samples deviate strongly from the true signal, they obey the same statistical properties, which is important for posterior analysis. The results indicate that the statistical properties (and ultimately the dynamics encoding field) are correctly reconstructed.

As a validation, in figures 5 and 6, we compare the
FIG. 5. **Top:** True Green’s function (left) and corresponding posterior mean (right). **Bottom:** Residual of the true Green’s function and the reconstruction (left) and a posterior sample for the Green’s function (right).

FIG. 6. Natural logarithmic spectrum of the true Green’s function (left) as well as the corresponding posterior mean (right).
reconstructed propagator as well as the corresponding spectrum with the underlying ground truth. The spectrum is comparable to the ground truth in regions with sufficient SNR while it levels out in regions where we have no information given via data. In addition, the reconstructed propagator also shows oscillations consistent with the true propagator. However we notice that modes that propagate “downwards” are reconstructed well while the weaker “upwards” propagating modes are not reconstructed due to the fact that they are below the noise level. In addition, we see that deviations in posterior samples of the propagator only occur within a “cone” and remain zero outside. This is due to the fact that we also reconstruct the maximal propagation speed of the process, which is reconstructed to be $c \approx 0.45 \ (\pm 0.08)$ enclosing the correct value of $c = 0.4$.

**C. Source detection**

Due to the fact that the prior breaks the phase degeneracy of the problem, the algorithm is also able to perform source detection in the excitation field, even in cases where the dynamics is unknown. To demonstrate this scenario we again generate a 1+1-dimensional mock example where in this case we assume that we are only able to measure the temporal evolution of the system at several locations. In particular we measure the temporal evolution at 50 randomly selected locations of the space under consideration. This results in $\approx 80\%$ of the discrete space-time being unobserved, as the resolution is the same as in the previous example. As before, we assume that the measurements are subject to additive Gaussian noise with $\sigma_n = 0.3$ and also assume the system to be at rest at $t = 0$. The resulting data is shown in the top-left panel of fig. The unknown excitations are inverse gamma distributed to model strong but sparse excitations. We infer those from measurements of the system at multiple locations together with the dynamic
response. The system used to generate the data exhibits damped traveling waves described by

\[ L = \partial_t^2 - c^2 \partial_x^2 + m^2 + \alpha \partial_t - \beta \partial_x , \]  
\[ (32) \]

with \((c, m, \alpha, \beta) = (1.2, 0.04, -0.013, 0.005)\). We again aim to reconstruct the dynamics encoding field, the propagation speed and the excitations from the data. From an information theoretical point of view this setup is very similar to the previous one since we only changed the measurement response to describe measurements of the temporal evolution at several locations, as well as the prior for excitations to be an inverse gamma prior. Consequently also the inference can be performed in the same way as before.

The setup as well as the reconstruction of the field evolving in space-time are shown in fig. 7. We see that the reconstruction recovers many sources and the corresponding propagation. The algorithm uses information from the response of strong excitations to reconstruct the Green’s function of the system. Due to the assumed homogeneity in space-time, this information helps to improve the overall reconstruction in other regions. The quality of the reconstruction of single excitations depends on the surrounding measurement scenario.

In figures 8 and 9 we depict the dynamic propagation as well as the spectrum and reconstructions. We can validate that the dynamical system was indeed reconstructed correctly, within uncertainties. We conclude that the task of source detection is possible even in cases where the underlying dynamics is unknown, as long as the assumptions of homogeneity, causality, and locality hold.

\[ \text{D. Non-linear evolution} \]

In the last example we aim to study the performance of the proposed method in an inconsistent setup. In particular we want to study the results of a reconstruction with data generated from a system which evolves non-linearly. Our inference algorithm is designed for linear autonomous systems, however we want to see what the resulting reconstruction extracts from data obtained from such systems.

1. Setup

To do so, we generate a solution of the diffusive Burgers equation, from which we generate data. The differential equation reads

\[ \partial_t \phi + \mu \ \partial_x^2 \phi + \alpha \ \phi \ \partial_x \phi = 0 , \]  
\[ (33) \]

where \(\mu\) and \(\alpha\) are constants.

We expect that the linear part of the dynamics will be captured by the differential operator we aim to infer, while non-linear interactions consequently have to be explained via consistent excitations. If we rewrite \(\phi\) as

\[ \phi_{t,x} = \tilde{\phi} + \delta \phi_{t,x} , \]  
\[ (34) \]

with

\[ \tilde{\phi} = \frac{1}{V} \int \phi_{t,x} \ dt \ dx , \]  
\[ (35) \]

where \(V\) is the space-time volume of the total space under consideration, we can rewrite eq. \[33\] as

\[ (\partial_t + \mu \ \partial_x^2 + \alpha \ \partial_x \phi) \delta \phi = -\alpha \ \delta \phi \ \partial_x \delta \phi \equiv \xi , \]  
\[ (36) \]

where we separated the linear from the non-linear evolution and identified the non-linear term with an excitation field \(\xi\).

Here we already notice a major difference between the expected result in this case compared to the previous examples. We see that both, the excitations as well as the dynamical operator depend on \(\phi\), indicating that we cannot expect the reconstructed dynamics of one particular \(\phi\) to generalize to a different solution, \(\phi'\), of the equation. By definition, the Green’s function of an autonomous system is the same for all solutions. Therefore, treating a Green’s function recovered for one particular \(\phi\) as a complete description leads to a wrong description of a system described via Burgers’ equation.

This is an important observation, since in many applications, the goal of the inference of the dynamical system is to use the posterior solution for the Green’s function as an abstract description of the system. Therefore, in the next section we will discuss how to identify whether or not it is possible to generalize the resulting dynamics to consistently produce different solutions of the observed dynamical system.

But first we have to validate if our expectations are met in an actual application. In figures 10 and 11 we show the mock solution as well as the reconstruction for \((\mu, \alpha) = (0.06, 1)\). The spectrum of the linear dynamics matches the effective linear evolution from eq. \[36\] and the excitation field recovers the corresponding non-linear term. In this sense the reconstruction of \(\phi\) is consistent with the data, but the reconstructed dynamical system does not fully describe the dynamical process. By definition, the method is only capable to infer the dynamic properties up to linear order. As we discussed above, the resulting linearized description is only valid for this particular solution and is no longer valid for a different solution of the Burgers equation.

2. Generalizing the reconstructed dynamics

The possibility of generalizing a reconstructed Green’s function to different realizations of the observed system, strongly depends on the completeness of the description. To be precise, if the reconstruction captures all aspects of the evolution of the system, it can actually be regarded
FIG. 8. **Top:** True Green's function of the process defined in eq. 32 (left) and corresponding posterior mean (right). **Bottom:** Residual of the true Green's function and the reconstruction (left) and a posterior sample for the Green's function (right).

FIG. 9. Natural logarithmic spectrum of the true Green's function (left) as well as the corresponding posterior mean (right).
FIG. 10. **Top:** Space-time evolution of the solution of Burgers' equation with an initial Gaussian profile at $t = 0$ (left) and the corresponding non-linear term as defined in eq. [36] (right). **Bottom:** Reconstruction of the solution (left) and the corresponding excitation field (right).

FIG. 11. Natural logarithmic spectrum of the effective linearized Green's function (left) for the Burgers equation (eq. [36]) and the spectrum of the corresponding reconstruction (right).
as a fundamental and complete description of the system rather than an effective one.

However, in an actual application we usually do not know a priori whether a linear autonomous description of the system is sufficient. Therefore we aim to use posterior information in order to make a statement about completeness. The posterior solution is fully characterized in terms of two constituents: the reconstructed Green’s function and the excitation field. As we have seen in the previous section, a non-linear interaction term results in a modified Green’s function and an excitation field that depends on the solution itself. The modification of the Green’s function is still a perfectly valid Green’s function in terms of our prior assumptions. Therefore it cannot provide information about the completeness of the description.

Consequently we have to rely on the recovered excitations. A priori the statistical properties of the excitation field are assumed to be known. Therefore, if the posterior solution of the excitations is a typical realization of the prior statistics, we may conclude that the recovered dynamical system is indeed a complete description of the data.

On the other hand, if this is not the case, the description might be incomplete because it is unlikely to reproduce the observed system via sampling from the prior of the excitations. At this point, it is not possible to decide whether the dynamic evolution of the system is not linear or autonomous, or if the prior assumptions on the excitations are wrong. Therefore, we cannot make a statement about the origin of the inconsistency, only that the description is likely to be incomplete, given the data.

As an illustrative example we can take a look at the posterior power spectra of the excitations to test for typicality. We notice that the power spectrum of the excitation field for the inconsistent setup (left panel of fig. 12), indicates that the posterior solution is correlated in space-time. In contrast, the power spectrum (right panel of fig. 12) of the posterior excitations from the consistent example of section V B is almost flat. This indicates that the posterior excitations are uncorrelated, as it was assumed a priori.

Besides power spectra, there are a variety of additional properties that can be used in order to test the posterior result on consistency. For example the assumption of homogeneity can be tested, only to name one. The Bayesian approach to this question would be to calculate the evidence for this model and compare it to the evidence of a different model. However, calculating the evidence is usually very difficult even approximately for fields.

Nevertheless, questions regarding the origin of deviations from prior assumptions remain a field of speculation without further information about the system. A more detailed discussion of this topic is beyond the scope of this work.

VI. CONCLUSION

In this work we considered the general problem of reconstructing the Green’s function as well as the excitation field of a linear autonomous SDE, in a blind setting, where the only external information given is the statistical properties of the excitation field as well as noisy and incomplete measurements of one corresponding solution. We show that the principles of locality and causality are sufficient to break the degeneracy in the likelihood between excitations and the Green’s function. Furthermore we construct a prior for the dynamic Green’s function which makes use of this information and derive the posterior distribution for a generic measurement scenario with Gaussian additive noise. As the posterior distribution of the problem at hand is intractable, we rely on a variational approximation in order to solve the inference problem.

To demonstrate its applicability, the proposed method is then applied to several consistent as well as inconsistent mock data examples. The results indicate that the method behaves as expected. The prior structure breaks the degeneracy of the likelihood, allowing for a unique reconstruction. In addition, the prior assumptions appear to be reasonable for physical applications.

In general, the proposed method aims to provide a solution to the task of inferring linear autonomous SDEs for systems that evolve in space and time. Hereby we rely on minimal prior assumptions, locality and causality, necessary to break the degeneracy between excitations and the Green’s function, which are physically motivated. This renders the method to be applicable in a
wide range of problems. One particular strength is the non-parametric formulation of the Green’s function. This becomes important in scenarios where physical models cannot provide a simple parametric description of evolution so far, to describe the Green’s function. In addition, a probabilistic description of excitations is sufficient for inference. Consequently the method is still applicable in cases where external influences cannot be described completely.

ACKNOWLEDGMENTS

We would like to thank Jakob Knollmüller, Philipp Arras and Margret Westerkamp for fruitful discussions, as well as Martin Reinecke for his contributions to NIFTy.

Appendix A: Light cone prior on a discretized space

As discussed in section [III E] the concept of causality in space-time results in a restriction of propagation within a light cone. In a continuous description, this restriction is realized via a convolution with a step function of the form

$$\Theta(t^2) = \Theta \left( t^2 - x^\dagger C^{-1} x \right) .$$  \hspace{1cm} (A1)

Due to the fact that our calculations are ultimately performed on a finite grid, this definition appears to be somewhat problematic, as it introduces boundary effects along the edges of the step function, when realized on a discretized space. In addition, if we aim to elevate \( C \) to be an unknown parameter of the problem that has to be inferred, gradient based methods are no longer applicable due to the fact that the gradient is zero almost everywhere in space-time (or not defined on the boundary). Therefore we seek to find a way to relax the sharp boundary introduced via the cone, without loosing its useful properties. To do so we borrow an idea from quantum field theory where it turns out that these sharp boundaries are “smeared” out when considering them on a quantum scale. In our case the “quantum” scale can be regarded as the resolution of the discrete representation of space-time, although this analogy is purely artificial.

To achieve this relaxation, consider the following quantity

$$\Delta = \sqrt{-t^2} = \sqrt{-(t^2 - x^\dagger C^{-1} x)} .$$  \hspace{1cm} (A2)

It has two useful properties: For causal (including time-like as well as light-like) points the real part of \( \Delta, \Re(\Delta) \), is zero as the square root is taken from a negative number. For non-causal (space-like) points the real part becomes positive. Furthermore, for fixed \( t \), \( \Re(\Delta) \) is asymptotically linear in \( x \). Therefore, if we consider a Gaussian in \( \Re(\Delta) \),

$$G(\Delta) \equiv \exp \left( -\frac{1}{2\sigma^2} \Re(\Delta)^2 \right) ,$$  \hspace{1cm} (A3)

we notice that this quantity remains one within the light cone, while it asymptotically falls off like a Gaussian in \( x \), for fixed \( t \). Here \( \sigma \) stands for an optional scaling parameter which controls the width of the Gaussian. In all applications of this paper it was set to a size of a few pixels.

[1] H. A. Kramers, Physica, 7, 284 (1940).
[2] H. Mori, Progress of Theoretical Physics, 33, 423 (1965).
[3] F. Black and M. Scholes, Journal of Political Economy 81, 637 (1973).
[4] A. N. Riseth and J. P. Taylor-King, arXiv e-prints, arXiv:1709.05153 (2017). [arXiv:1709.05153 [math.ST]]
[5] P. Frank, T. Steininger, and T. A. Enßlin, Phys. Rev. E 96, 052104 (2017).
[6] N. Oppermann, M. Selig, M. R. Bell, and T. A. Enßlin, Phys. Rev. E 87, 032136 (2013). [arXiv:1210.6866 [astro-ph.IM]]
[7] T. Kailath, A. H. Sayed, and B. Hassibi, Linear Estimation (Pearson, 2000).
[8] N. Wiener, Extrapolation, interpolation, and smoothing of stationary time series, with engineering applications.
[9] T. A. Enßlin, M. Frommert, and F. S. Kitaura, Phys. Rev. D 80, 105005 (2009). [arXiv:0806.3474 [astro-ph]].
[10] T. A. Enßlin, arXiv e-prints (2018). [arXiv:1804.03350]
[11] J. Knollmüller, T. Steininger, and T. A. Enßlin, arXiv e-prints, arXiv:1711.02955 (2017). [arXiv:1711.02955 [stat.ME]].
[12] J. Knollmüller and T. A. Enßlin, arXiv e-prints, arXiv:1812.04403 (2018). [arXiv:1812.04403 [astro-ph]].
[13] M. Selig, M. R. Bell, H. Junklewitz, N. Oppermann, M. Reinecke, M. Greiner, C. Pachajoa, and T. A. Enßlin, A&A 554, A26 (2013). [arXiv:1301.4499 [astro-ph.IM]].
[14] T. Steininger, J. Dixit, P. Frank, M. Greiner, S. Hutschenreuter, J. Knollmüller, R. Leike, N. Porqueres, D. Pumpe, M. Reinecke, M. Šraml, C. Varady, and T. Enßlin, arXiv e-prints, arXiv:1708.01073 (2017). [arXiv:1708.01073 [astro-ph.IM]].
[15] http://ift.pages.mpcdf.de/NIFTy/, to be released on 01.02.2019.
[16] J. Knollmüller and T. Enßlin, in preparation.