An Additive Approximation to Multiplicative Noise

R. Nicholson1 · J. P. Kaipio2

Received: 13 March 2019 / Accepted: 21 July 2020 / Published online: 17 August 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Multiplicative noise models are often used instead of additive noise models in cases in which the noise variance depends on the state. Furthermore, when Poisson distributions with relatively small counts are approximated with normal distributions, multiplicative noise approximations are straightforward to implement. There are a number of limitations in the existing approaches to deal with multiplicative errors, such as positivity of the multiplicative noise term. The focus in this paper is on large dimensional (inverse) problems for which sampling-type approaches have too high computational complexity. In this paper, we propose an alternative approach utilising the Bayesian framework to carry out approximative marginalisation over the multiplicative error by embedding the statistics in an additive error term. The Bayesian framework allows the statistics of the resulting additive error term to be found based on the statistics of the other unknowns. As an example, we consider a deconvolution problem on random fields with different statistics of the multiplicative noise. Furthermore, the approach allows for correlated multiplicative noise. We show that the proposed approach provides feasible error estimates in the sense that the posterior models support the actual image.

Keywords Multiplicative noise · Additive approximation · Pre-marginalisation

1 Introduction

A ubiquitous problem in science and engineering is to infer some parameter of interest, say $x \in \mathbb{R}^n$, given noisy indirect measurements $y \in \mathbb{R}^m$. Suppose the parameter and measurements are linked by a parameter-to-observable map $f : \mathbb{R}^n \times \mathbb{R}^P \times \mathbb{R}^q \rightarrow \mathbb{R}^m$, so that we can write

$$y = f(x, n, \eta),$$

(1.1)

where $n \in \mathbb{R}^P$ and $\eta \in \mathbb{R}^q$ denote uninteresting random variables, which can be interpreted as noise.

Here, we consider the inference problem in the Bayesian framework [18,36,37], which naturally allows for the incorporation of uncertainties and prior knowledge, and results in a posterior distribution. In this framework, a natural first task would then be to marginalise over the uninteresting variables. However, the marginalisation process depends explicitly on how the noise is modelled. The most common model for $f(x, n, \eta)$ is the additive error model [18,37]

$$y = A(x) + \eta,$$

(1.2)

where the mapping $A : x \mapsto y$ is referred to as the forward map (problem). However, in several imaging modalities including optical coherence tomography (OCT) [41,42], ultrasound [5,27], synthetic aperture radar (SAR) imaging [12,40], and electrical impedance tomography (EIT) [2,43], noise can be proportional to the data. Multiplicative noise is also common in systems and control theory [3]. In such a case, we have

$$y = n \odot A(x),$$

(1.3)

where $\odot$ denotes component-wise (Hadamard) product. Such a set-up is usually referred to as the multiplicative noise model.

In many cases, there may simultaneously be multiplicative and additive noise present, see, for example, [11,13,18,22,30]. In such case, we can write

$$y = n \odot A(x) + \eta.$$

(1.4)
However, often the effects of the additive errors $\eta$ have been assumed to be small (compared to the effects of the multiplicative noise, $n$) and are systematically neglected [11,22,25,35]. Furthermore, the components of the multiplicative noise are essentially always assumed to be mutually independent. In the current paper, however, we retain the additive error term and also consider correlated multiplicative noise.

In this paper, we take a model discrepancy style approach to transform Eq. (1.4) to the form of Eq. (1.2), but with a modified additive error term, which reflects both the original additive errors and multiplicative errors. The updated error term is then approximated as Gaussian and subsequently marginalised over, resulting in a convex likelihood. Furthermore, if the prior distribution is Gaussian and the forward map is linear, then the computation of the associated MAP estimate and posterior covariance matrix can be computed analytically using only linear algebra.

The paper is organised as follows. In Sect. 2, we review the marginalisation of noise terms in the Bayesian framework. In Sect. 3, we give a brief review of the methods used to deal with multiplicative noise. Section 4 outlines the approximation of noise statistics and the subsequent marginalisation, which is sometimes referred to as the Bayesian approximation error (BAE) approach [17–19]. The multiplicative noise term is not assumed to be uncorrelated. In Sect. 5, we consider a deconvolution example with different distributions for the multiplication noise, including correlated multiplicative noise.

## 2 Exact Marginalisation Over Additive and Multiplicative Noise Terms

In this paper, we assume that the noise terms $n$ and $\eta$ and the parameter of interest $x$ are independent of each other. Thus, the joint model of the noise terms and the parameter can be stated as $\pi(x, n, \eta) = \pi(x)\pi_n(n)\pi_\eta(\eta)$. For now, inline with the literature [11,12,34,35,44], we assume each component of the multiplicative noise is independent and identically distributed (iid), allowing us to write $\pi_n(n) = \prod_{i=1}^m \pi_{ni}(n_i)$, though this assumption is later dropped (see Sect. 5.3).

The likelihood is obtained formally by marginalisation

$$
\pi(y|x) = \int \int \pi(y|x, n, \eta)\pi_n(n)\pi_\eta(\eta) \, dn \, d\eta
$$

$$
\pi(y|x) = \int \int \delta(y - n \odot A(x) - \eta)\pi_n(n)\pi_\eta(\eta) \, dn \, d\eta.
$$

(2.1)

where $\delta(\cdot)$ is the Dirac distribution, which assigns all mass at the origin. We now look at the three individual cases of interest.

In the purely additive noise model, we set $\pi(n) = \delta(n - 1) = \prod_{i=1}^m \delta(n_i - 1)$. Thus, (2.1) can be written as

\[
\pi(y|x) = \int \int \delta(y - n \odot A(x) - \eta)\delta(n - 1) \, dn \pi_\eta(\eta) \, d\eta
\]

\[
= \int \delta(y - A(x) - \eta)\pi_\eta(\eta) \, d\eta = \pi_\eta(y - A(x)).
\]

(2.2)

In the purely multiplicative noise model, we set $\pi(\eta) = \delta(\eta) = \prod_{i=1}^m \delta(\eta_i)$, then (2.1) can be written as

\[
\pi(y|x) = \int \int \delta(y - n \odot A(x) - \eta)\delta(\eta) \, dn \pi_n(n) \, d\eta
\]

\[
= \int \delta(y - n \odot A(x))\pi_n(n) \, dn
\]

\[
= \prod_{i=1}^m \left( \frac{1}{|A_i(x)|} \pi_{ni} \left( \frac{y_i}{A_i(x)} \right) \right),
\]

(2.3)

where $A_i(x)$ is the $i$th component of $A(x)$.

In the case of simultaneous multiplicative and additive noise terms, the integrations in (2.1) can be carried out in either order resulting in either

\[
\pi(y|x) = \int \int \delta(y - n \odot A(x) - \eta)\pi_n(n)\pi_\eta(\eta) \, dn \, d\eta
\]

\[
= \int \pi_n(n)\pi_\eta(y - n \odot A(x)) \, dn,
\]

(2.4a)

or

\[
\pi(y|x) = \int \int \delta(y - n \odot A(x) - \eta)\pi_n(n)\pi_\eta(\eta) \, d\eta \, dn
\]

\[
= \prod_{i=1}^m \left( \frac{1}{|A_i(x)|} \right) \int \prod_{i=1}^m \left( \pi_{ni} \left( \frac{y_i - \eta_i}{A_i(x)} \right) \right) \pi_\eta(\eta) \, d\eta.
\]

(2.4b)

It is shown in “Appendix A” that (2.4a) and (2.4b) are equal when $n$ and $\eta$ are absolutely continuous random variables. Independently of which form of the likelihood is favoured, the integration cannot generally be computed analytically.

## 3 Approaches to Handle Multiplicative Noise Models

For the remainder of the paper, we will consider linear forward models $A(x) = Ax$, as is the case in deblurring (setting $A = I$ is often referred to as denoising). Furthermore, since...
the focus of the present paper is on inverse problems and since some approaches depend directly on properties of the unknown (such as positivity), we refer directly to posterior models. Moreover, since the proposed approach is targeted at relatively large dimensional problems, we will consider the computation of MAP estimates and local Gaussian approximations to the posteriors only.

There are several approaches documented in the literature for dealing with multiplicative noise. Many of the techniques are framed in the context of denoising and are based on the use of the total variation (TV) prior and assume positivity of the unknown, \( x \), see, for example, [1,11,30,34,35,44]. In the present paper, we do away with these assumptions, but note that the proposed approach results in only a reformulated likelihood while leaving the prior unchanged. Thus, inclusion of various prior models could therefore naturally be incorporated.x

We now briefly outline some of the most commonly used models.

The Log Model (Multiplicative Noise Only): The most common of these techniques is to apply the logarithm transform method, resulting in a problem of the form of (1.2), see, for example, [11,14]. However, there are some drawbacks to applying the logarithm transform method. Firstly, if any of the components of the data, \( y \) or the model prediction, \( Ax \), are negative, the method fails. On the other hand, if the additive error is retained, as in Eq. (1.4), the logarithm transform is of little use. Furthermore, it has been noted, that by using this method one cannot always directly apply standard additive noise removal algorithms and results are often unsatisfactory [1]. That being said, for a general prior, \( \pi_s(x) \), the approach leads to a MAP estimate of the form

\[
x_{\text{MAP}} = \arg \max_{x \in \mathbb{R}^n} \pi_{\xi}(\log(y) - \log(Ax)) \pi_s(x),
\]

where \( \pi_{\xi} \) is the density of \( \xi = \log(n) \). Iterative methods must then be used to solve for \( x_{\text{MAP}} \). The basic idea of transforming multiplicative noise to additive noise is, in principle, similar to the procedure we propose in the current paper. However, in this paper, neither the measurements nor the forward model is transformed, which is particularly convenient in the case of linear forward models.

The AA Model: The so-called AA model was derived in [1] for the MAP estimate under the assumption that the multiplicative noise follows a Gamma distribution, and that the data, \( y \), and the unknown, \( x \), are positive. The MAP estimate, assuming \( x \) has a Gibbs prior distribution with potential \( \phi(x) \), is then of the form

\[
x_{\text{MAP}} = \arg \min_{x \in \mathbb{R}^n} \sum_{i=1}^n \left( L \left( \log(A_i x) + \frac{y_i}{A_i} \right) + \phi(x) \right),
\]

(3.2)

where \( A_i \) denotes the \( i \)-th row of \( A \). The computation of the MAP estimate in this case also requires iterative methods even when the prior on \( x \) is Gaussian. Furthermore, the likelihood potential is not always strictly convex, although the existence of a minimiser was proven in [1].

The Separable Model: The separable model was introduced in [15] and takes into account both additive and multiplicative noise. Furthermore, for several different multiplicative noise models, closed-form functionals are derived for the associated MAP estimate. Here, we give a brief outline of how the posterior is found. In accordance with [15], the separable model takes the form

\[
y = f(x, n, \eta) = n \odot (Ax + \eta) ,
\]

(3.3)

where \( y \) is assumed to be positive. The separable model relies on the introduction of an intermediate variable,

\[
u = Ax + \eta ,
\]

(3.4)

which is also assumed to be positive. The full posterior can then be written as

\[
\begin{align*}
p_{\text{post}}(u, x | y) & \propto \pi(y | u, x) \pi(u, x) \\
& = \pi(y | u) \pi(u | x) \pi_s(x) .
\end{align*}
\]

(3.5)

The conditional densities in (3.5) can be derived in a similar manner to how the likelihood densities were dealt with in Sect. 2. Firstly,

\[
\pi(y | u) = \int \delta(y - n \odot u) \pi_n(n) \, dn = \prod_{i=1}^m \left( \frac{1}{|u_i|} \pi_{u_i} \left( \frac{y_i}{u_{i}} \right) \right) ,
\]

(3.6)

and secondly,

\[
\pi(u | x) = \int \delta(\eta - u - Ax) \pi_\eta(\eta) \, d\eta = \pi_\eta(u - Ax) .
\]

(3.7)

This results in a posterior of the form

\[
\pi(u, x | y) = \prod_{i=1}^m \left( \frac{1}{|u_i|} \pi_{u_i} \left( \frac{y_i}{u_{i}} \right) \right) \pi_\eta(u - Ax) \pi_s(x) .
\]

(3.8)
The MAP estimates \((u_{\text{MAP}}, x_{\text{MAP}})\) are given in closed form for several prior densities on the multiplicative noise in [15]. The main drawbacks to this method are the required positivity of both the data, \(y\), and the intermediate variable, \(u\), as well as the lack of convexity for the functionals which need to be minimised in order to calculate the MAP. Furthermore, computation of the MAP estimate again requires the use of iterative methods independent of the choice of prior model and will need to be carried out for the number of primary unknowns and the number of measurements so as to estimate \(x\) and \(u\).

Other methods for dealing with multiplicative noise in the denoising context include filtering-type approaches such as those discussed in [13] and the construction and use of similarity measures [39]. For filtering-type methods, the problem is framed in the so-called space-time formalism. On the other hand, approaches using similarity measures usually set values of the restored image to some weighted mean of the surrounding pixels, where the weights depend on the similarity of the pixels.

### 4 Approximate Marginalisation of Multiplicative Noise

In this paper, we carry out approximative marginalisation over both the additive and multiplicative noise terms prior to conditioning on the data. In the inverse problems literature, this approach is referred to as the Bayesian approximation error (BAE) approach since the approximative marginalisation is carried out over the prior distribution. The BAE approach was introduced in [18,19] to take into account the discrepancy between accurate and reduced-order models. Since then, the approach has been extended, for example, to account for errors and uncertainties related to uninteresting distributed parameters in PDE’s [21], errors in the geometry of the domain [29], unknown boundary data [23], approximation of the (physical) forward map [38], and state estimation problems [16,24,26]. For a more general discussion of the approach and a more extended list of extensions, see [17]. Below, we adapt the approach to the context of multiplicative noise.

The goal is to embed the additive and multiplicative noise terms in an approximate Gaussian additive error only model. Without such an approximation, the updated likelihood is convex, and in the case of a linear forward model and Gaussian prior, the computation of the approximate MAP estimate and the approximate posterior covariance reduces to linear algebra. This is in contrast to standard approaches for dealing with nonconvexity issues, which as alluded to can suffer from nonconvexity issues, and essentially always require iterative methods [1,11,14,15,30,34].

With the present observation model, we can write

\[
y = n \odot Ax + \eta = Ax + (n - 1) \odot Ax + \eta = Ax + \epsilon + \eta = Ax + \epsilon,
\]

which is an alternative and exact additive error model to use in place of (1.4). Formal marginalisation over \(\epsilon\) would then yield the likelihood model \(\pi(y \mid x) = \pi_{e \mid x}(y - Ax \mid x)\), the computation of which is, however, not generally possible analytically. In the BAE approach, at this stage, one (usually) makes the Gaussian approximation \(\pi_{e \mid x}(\epsilon) = N(e_{\mid x}, \Gamma_{e \mid x})\), see, for example, [17–19]. A small amount of work has been carried out on retaining the full density of the errors, but these rely on sampling-based approaches (even in linear cases) [6,7]. Furthermore, they have only been implemented in the case of additive noise. In theory, however, the full density of the approximation errors could be calculated as the product density of \(p \odot q\) for \(p = n - 1\) and \(q = Ax\), see, for example, [31].

By taking the Gaussian approximation to \(\pi_{e \mid x}(\epsilon \mid x)\), we only require the mean, \(E(\epsilon \mid x) = e_{\mid x}\), and the covariance, \(\Gamma_{e \mid x}\). For the marginal mean \(E(\epsilon) = e_s\), we have

\[
e_s = \eta_s + (n_s - 1) \odot Ax_s.
\]
where \( L_{e|x}^T L_{e|x} = \Gamma_{e|x}^{-1} \) and is typically found using the Cholesky decomposition \([9]\). In the case of iid additive noise, iid multiplicative noise, and that \( x \) and \( \eta \) are independent (which are the standard assumptions), we would have

\[
\begin{align*}
\Gamma_{ee} &= \sigma_n^2 I + \sigma_x^2 \text{diag} \left( A \Gamma_{xx} A^T \right) \\
\Gamma_{ex} &= 0,
\end{align*}
\]

(4.5) (4.6)

where \( \sigma_n \) and \( \sigma_x \) denote the standard deviation of the additive noise and multiplicative noise, respectively, and thus \( \Gamma_{e|x} = \Gamma_{ee} \). Letting \( L_e^T L_e = \Gamma_{ee}^{-1} \), we can rewrite the approximate likelihood model as

\[
\pi(y \mid x) \propto \exp \left( -\frac{1}{2} \| L_e (y - Ax) \|_2^2 \right).
\]

(4.7)

Note that the structure of the covariance \( \Gamma_{ee} \) is, in the general case, nontrivial and depends also on the prior covariance \( \Gamma_{xx} \), which is typical in BAE-type approaches. As far as the authors are aware, correlated multiplicative noise has not previously been considered in the literature.

5 Application to Deblurring

We consider an image deblurring (deconvolution) example with three different multiplicative noise statistics: normal, Gamma, and uniform distributions. The image used assumes both positive and negative values. For each case, additive noise with standard deviation corresponding to 1% of the range of the noiseless observations. We also consider the same example with normal multiplicative noise that is spatially correlated. Since the focus in this paper is on the multiplicative noise, we take the image to be uncorrelated with the additive noise component.

5.1 The Target, the Observations, the Prior Model, and the Additive Noise Model

For all examples, we specify a 50 \( \times \) 50 pixel target image shown in Fig. 1. We blur the image with a symmetric Gaussian blurring kernel

\[
\mathcal{K}(s_1, s_2) = \frac{1}{2\pi \kappa^2} \exp \left( -\frac{s_1^2 + s_2^2}{2\kappa^2} \right),
\]

(5.1)

with \( \kappa = 5 \) also shown in Fig. 1. Both the image and the kernel are taken to be piecewise constant in a grid with rectangular elements. We take the forward operator to be the circulant convolution operator \( \mathcal{K} \) [8], so that we can write

\[
y = n \otimes (K \ast x) + \eta = n \otimes Ax + \eta,
\]

(5.2)

where \( K \) is the circulant realisation of the kernel \( \mathcal{K} \) and, further, \( A \) is the realisation of \( K \) in matrix form. The blurred (noiseless) image \( y = Ax \) is also shown in Fig. 1. The observations with the three different multiplicative noise models are shown in Fig. 4.

In this paper, we employ a normal prior model \( x \sim \mathcal{N}(x_0, \Gamma_{xx}) \). The mean of \( x \) is set to be spatially homogeneous, i.e. \( E(x) = x_0 = x_n 1 \). For the prior covariance matrix, we employ so-called PDE-based covariance matrices \([4,28,36]\). The PDE-based covariance matrices used here are based on the finite element method (FEM). Specifically, we create a finite element mesh consisting of triangular elements, by identifying the center of each pixel as a finite element node, and then (Delaunay) triangulating the points. We denote by \( \Omega \) the finite element domain and use the standard Lagrange piecewise linear nodal basis functions, which we denote by \( \phi_i \).

The prior covariance matrix is set as

\[
\Gamma_{xx} = (c_1 (c_2 G + M))^{-2} = \left( L_e^T L_e \right)^{-1},
\]

(5.3)

where the matrices \( G \) and \( M \) are the stiffness and mass matrices, respectively, i.e.

\[
G_{ij} = \int_\Omega \nabla \phi_i \cdot \nabla \phi_j \, ds,
\]

\[
M_{ij} = \int_\Omega \phi_i \phi_j \, ds, \quad i, j = 1, 2, \ldots, n.
\]

(5.4)

The parameter \( c_1 \) is inversely proportional to the prior standard deviation, while the product \( c_2 c_4 \) controls the correlation length, see, for example, \([4,10]\). This choice of prior covariance matrix implicitly defines homogeneous Neumann boundary conditions on the prior covariance. However, other boundary conditions can naturally be implemented, see, for example, \([10,20,32]\). The matrix square root \( L_e \) can also be interpreted as the whitening operator of the random field \([33]\).

The parameters are set here as \( c_1 = 10^{-1} \) and \( c_2 = 20 \) so that the range of \( x \) and the correlation length are approximately consistent with the structure of the target image, we also set \( x_n = 0 \). See Fig. 2 for the covariance function and two draws from the prior model.

Finally, the additive noise model is \( \eta \sim \mathcal{N}(0, \sigma_n^2 I) \). A correlated additive noise model is straightforward to handle as shown in Sect. 4.

5.2 Reconstructions with Spatially Uncorrelated Multiplicative Noise

Without loss of generality, we set \( E(n) = 1 \) for all cases, as is customary \([1,25,34]\). Furthermore, in this section, we take the components of \( n \) to be iid so that \( \Gamma_{nn} = \sigma_n^2 I \). We consider three different distributions for the multiplicative noise, \( n \),
Fig. 1 Left: the target image $x_{\text{true}}$. Centre: the Gaussian convolution kernel $K$ centred at the centre of the image. Right: the blurred noiseless image $K \ast x$

Fig. 2 Left: The correlation function (scaled so that the maximum is the marginal variance) induced by the PDE-based model with $c_1 = 10^{-1}$ and $c_2 = 20$. Centre and right: two draws from the prior model and scale them so that the variances coincide. Furthermore, in two cases, the probability $P(n_i < 0)$ does not vanish.

The first model for $n$ is the iid Gamma distribution, which has been the most common model for multiplicative noise [1,11,34], and is written

$$n_i \sim \Gamma(\alpha, \beta), \quad i = 1, 2, \ldots, m,$$

(5.5)

where $\alpha$ and $\beta$ are the shape and scale parameters, respectively. For $n$ such that $\mathbb{E}(n) = 1$, we can also write

$$n_i \sim \Gamma\left(L, \frac{1}{L}\right), \quad i = 1, 2, \ldots, m.$$

(5.6)

We set $L = 1$ so that $\text{var}(n_i) = 1$. For the Gamma distribution, $P(n_i < 0) = 0$.

The second model for $n$ which is more seldom considered is the iid normal model

$$n_i \sim \mathcal{N}(1, \sigma^2), \quad i = 1, 2, \ldots, m,$$

(5.7)

where throughout the literature $\sigma \leq 0.2$ is referred to as tiny noise, see, for example, [1]. The assumption of tiny noise is done in an attempt to avoid multiplicative noise terms becoming negative, as discussed in Sect. 3. In the approach proposed in this paper, we do not need to make such an assumption and we set, again, $\sigma = 1$ which results in $P(n_i < 0) \approx 0.15$.

As the third model, we consider multiplicative noise with iid uniform distribution,

$$n_i \sim \mathcal{U}(1 - \nu, 1 + \nu), \quad i = 1, 2, \ldots, m$$

(5.8)

and we set $\nu = \sqrt{3}$ so that, again, $\text{var}(n_i) = 1$ and which results in $P(n_i < 0) \approx 0.21$.

Draws from the three multiplicative noise models are shown in Fig. 3.

The reconstructions computed using the proposed approximation, denoted $x_{\text{MAP}}^{\gamma}$, $x_{\text{MAP}}^{\text{normal}}$, and $x_{\text{MAP}}^{\text{uniform}}$ are shown in Fig. 5. Furthermore, in the bottom row of Fig. 5, we show the estimates and posterior confidence intervals along the cross section shown in the images in the top row. We see that embedding the multiplicative noise into the additive error leads to feasible results in the sense that the actual target is supported by the approximative MAP $\pm 3\sqrt{\mathbb{E}(y(k)|x)}$ intervals. It is worth pointing out that in the case of normal and uniform multiplicative noise distributions, $n$ exhibits negative samples. Clearly, this does not constitute a problem for the proposed approach. The feasibility of the posterior estimates is similar with all three distributions of the multiplicative noise. The fact that the estimates obtained here are fairly smooth in comparison with the true image is due to the use of a Gaussian smoothness prior. To reconstruct the sharp edges, one could employ a TV type prior [1]. Even with a normal approximation for the posterior, this would result in the need for iterative methods to compute the MAP estimate.

5.3 Reconstructions with Spatially Correlated Multiplicative Noise

The derivations of the existing methods to handle multiplicative noise are largely based on the assumed iid property. In the proposed approach, such an assumption is not required as
Fig. 3 Draws from the different iid multiplicative noise distributions. Left: Gamma distribution $\Gamma(1, 1)$. Centre: normal distribution $\mathcal{N}(1, I)$. Right: uniform distribution $\mathcal{U}(1 - \sqrt{3}, 1 + \sqrt{3})$.

Fig. 4 Data corrupted by uncorrelated multiplicative noise generated from left: gamma, centre: normal, and right: uniform distributions.

Fig. 5 Top row: the MAP estimates attained by using the BAE approach with different iid multiplicative noise models. Left: Gamma, centre: normal, and right: uniform noise models. Bottom row: cross sections of the actual target and reconstructions with approximate MAP $\pm 3\sqrt{\Gamma_{e|x}(k, \bar{k})}$ intervals along the lines in the top row reconstructions.

indicated by the approximate joint covariance $\Gamma_{e,x}$ in Sect. 4. With deblurring problems such as the present example, it is clear that when the spatial correlation structure of the multiplicative noise gets more complicated, we can expect the actual estimation errors to increase. This can be expected, in particular, with noise distributions which have relatively long and positive correlation.

In this section, we consider Gaussian multiplicative noise with different spatial correlation decay rates. The traces of the multiplicative noise covariances are the same as in the cases of spatially uncorrelated noise. Furthermore, the variance of the (spatially uncorrelated) additive noise is as in the previous cases. The respective correlation functions and draws from these distributions are shown in Fig. 6. The respective observations are shown in Fig. 7.

The approximate MAP estimates and the posterior $\pm 3$ STD intervals are shown in Fig. 8. The estimates are, again, feasible with respect to the posterior error intervals. The error estimates are larger than in the case of spatially uncorrelated multiplicative noise which was expected. As was also
Fig. 6  Top row: spatial correlation models for the multiplicative noise with different spatial decay rates. Bottom row: Draws from the respective models.

Fig. 7  The observations with the three different spatially correlated multiplicative noise models shown in Fig. 6.

Fig. 8  Top row: the MAP estimates attained by using the BAE approach with different spatially correlated multiplicative noise models shown in Fig. 6. Left: Gamma, centre: normal, and right: uniform noise models. Bottom row: cross sections of the actual target and reconstructions with approximate MAP ±3√Γx[y](k, k) intervals along the lines in the top row reconstructions.
expected, the error estimates increase with decreasing decay rate of spatial correlation of the multiplicative noise.

6 Conclusion

In this paper, we proposed an approach to approximate (linear) inverse problems corrupted by both additive and multiplicative noise with an additive noise model. The approximate additive noise model is constructed by incorporating the discrepancy between the model predictions of the original and the approximate model, which is referred to as the Bayesian approximation error (BAE) approach. The resulting additive noise term is then approximated as a Gaussian random variable. The covariance of this term is nontrivial and depends on the prior covariance. For linear forward problems, the computation of the approximate MAP estimate does not suffer from convexity-related problems and can be easily computed if the prior distribution is Gaussian.

The approach does not need the multiplicative noise to be uncorrelated. The results in this paper are, however, based on the primary unknown and the multiplicative noise being independent. As such, the independent of the primary unknown and the multiplicative noise is, however, not essential for the proposed approach. In such a case, the computation of the related joint covariance of the modified additive noise and the primary unknown involves rather tedious mappings of general fourth-order statistics (kurtosis).

We considered numerical examples with different multiplicative noise distributions related to an image processing deconvolution problem with both additive and multiplicative noise. The results show that the approximation is feasible in the sense that the posterior error estimates mostly support the actual target image. Furthermore, the results were feasible also when the multiplicative noise was highly spatially correlated.

Acknowledgements The authors would like to thank Tom ter Elst and Oliver Maclaren for their insights and also the anonymous reviewers for their helpful and constructive comments that greatly contributed to improving the final version of this paper.

A Equality of the Likelihoods

Here, we provide a formal proof that the two forms of the likelihood, given in (2.4a) and (2.4b), computed under simultaneous multiplicative and additive errors are equal. The assumptions used here are slightly more restrictive than necessary, but cover most applications and result in the proof being straightforward.

Theorem A.1 Let \( x \in \mathbb{R}^n, y, n, \eta \in \mathbb{R}^m \) and \( A : \mathbb{R}^n \rightarrow \mathbb{R}^m \) be continuous. Furthermore, assume \( n \) and \( \eta \) are absolutely continuous random variables. Then,

\[
\int \int \delta(y - n \odot A(x) - \eta) \, \pi_n(n) \, dn \, \pi_\eta(\eta) \, d\eta = \int \int \delta(y - n \odot A(x) - \eta) \, \pi_\eta(\eta) \, d\eta \, \pi_n(n) \, dn. \tag{A.1}
\]

Proof Firstly, as \( x \) is fixed in terms of the integrals in (A.1), define \( a = A(x) \). Furthermore, we will \( \odot \) to denote component-wise division.

Now, let \( \mathcal{I} \) and \( \mathcal{J} \) be subsets of the set \( \mathcal{S} := \{1, 2, \ldots, m\} \) with \( \mathcal{I} \cup \mathcal{J} = \mathcal{S} \) such that \( a_i = 0 \) for each \( i \in \mathcal{I} \) and \( a_i \neq 0 \) for each \( i \in \mathcal{J} \). Without loss of generality, we assume \( \mathcal{I} = \{1, 2, \ldots, p\} \) and \( \mathcal{J} = \{p+1, p+2, \ldots, m\} \) (note that \( p = m \) or \( p = 0 \), i.e. \( \mathcal{I} = \mathcal{S} \) and \( \mathcal{J} = \emptyset \) or \( \mathcal{I} = \emptyset \) and \( \mathcal{J} = \mathcal{S} \), respectively, are special cases of this). Then, taking \( dn = dn_{\mathcal{I}} \, dn_{\mathcal{J}} \) and \( d\eta = d\eta_{\mathcal{I}} \, d\eta_{\mathcal{J}} \), we can rewrite the left-hand side of Eq. (A.1) as

\[
\mathcal{L} = \int_{\mathbb{R}^{m-p}} \int_{\mathbb{R}^{m-p}} \delta(y - n \odot a - \eta) \, \pi_n(n) \, dn \, \pi_\eta(\eta) \, d\eta
\]

\[
= \int_{\mathbb{R}^{m-p}} \int_{\mathbb{R}^{m-p}} \int_{\mathbb{R}^p} \int_{\mathbb{R}^p} \delta_I(y_{\mathcal{I}} - \eta_{\mathcal{I}}) \times \delta_{\mathcal{J}}(y_{\mathcal{J}} - n_{\mathcal{J}} \odot a_{\mathcal{J}} - \eta_{\mathcal{J}}) \, \pi_n(n) \, dn_{\mathcal{I}} \, \pi_\eta(\eta) \, d\eta_{\mathcal{I}} \, d\eta_{\mathcal{J}}
\]

\[
= \int_{\mathbb{R}^{m-p}} \int_{\mathbb{R}^p} \delta_I(y_{\mathcal{I}} - \eta_{\mathcal{I}}) \int_{\mathbb{R}^{m-p}} \int_{\mathbb{R}^p} \delta_{\mathcal{J}}(y_{\mathcal{J}} - n_{\mathcal{J}} \odot a_{\mathcal{J}} - \eta_{\mathcal{J}}) \times \pi_n(n) \, dn_{\mathcal{I}} \, \pi_\eta(\eta) \, d\eta_{\mathcal{I}} \, d\eta_{\mathcal{J}}
\]

\[
= \int_{\mathbb{R}^{m-p}} \int_{\mathbb{R}^p} \delta_I(y_{\mathcal{I}} - \eta_{\mathcal{I}}) \int_{\mathbb{R}^{m-p}} \delta_{\mathcal{J}}(y_{\mathcal{J}} - n_{\mathcal{J}} \odot a_{\mathcal{J}} - \eta_{\mathcal{J}}) \times \pi_n(n) \, dn_{\mathcal{I}} \, \pi_\eta(\eta) \, d\eta_{\mathcal{I}} \, d\eta_{\mathcal{J}}
\]

\[
= \left( \prod_{i=p+1}^m |a_i|^{-1} \right) \int_{\mathbb{R}^{m-p}} \int_{\mathbb{R}^p} \delta_I(y_{\mathcal{I}} - \eta_{\mathcal{I}}) \times \pi_{n_{\mathcal{J}}}(y_{\mathcal{J}} - \eta_{\mathcal{J}} \odot a_{\mathcal{J}}) \, \pi_\eta(\eta) \, d\eta_{\mathcal{I}} \, d\eta_{\mathcal{J}}
\]

\[
= \left( \prod_{i=p+1}^m |a_i|^{-1} \right) \int_{\mathbb{R}^{m-p}} \pi_{n_{\mathcal{J}}}(y_{\mathcal{J}} - \eta_{\mathcal{J}} \odot a_{\mathcal{J}}) \times \pi_\eta(y_{\mathcal{I}} \odot a_{\mathcal{I}}) \, dn_{\mathcal{I}} \, d\eta_{\mathcal{J}}
\]

where

\[ \pi_{n_{\mathcal{J}}}(n_{\mathcal{J}}) = \int_{\mathbb{R}^p} \pi(n) \, dn_{\mathcal{I}} \]
is the marginal distribution. On the other hand, we can rewrite the right-hand integral of Eq. (A.1) as

\[ R = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \delta(y - n \otimes a - \eta) \pi_\eta(\eta) d\eta \pi_n(n) dn \]

\[ = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \delta(y_J - \eta_J) \times \delta(y - n_J \otimes a_J - \eta_J) \pi_\eta(\eta) d\eta_J d\eta_J \pi_n(n) dn_J d\eta_J \]

\[ = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \delta(y_J - n_J \otimes a_J - \eta_J) \times \pi_n(y_J, \eta_J) d\eta_J \pi_n(n) dn_J d\eta_J \]

\[ = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \delta(y_J - n_J \otimes a_J) \pi_n(n) dn_J d\eta_J \]

\[ = \int_{\mathbb{R}^m} \pi_n(y_J, a_J) \pi_n(n) dn_J. \]

Finally, by introducing the change of variables \( z_J = y_J - n_J \otimes a_J \), we have \( n_J = (y_J - z_J) \otimes a_J \), and

\[ dn_J = \left( \prod_{l=p+1}^m \delta_k^{-1} \right) dz_J, \]

and hence

\[ R = \int_{\mathbb{R}^m} \pi_n(y_J, z_J) \pi_n(n) dn_J \]

\[ = \left( \prod_{l=p+1}^m |\delta_k|^{-1} \right) \int_{\mathbb{R}^m} \pi_n(y_J, z_J) \pi_n((y_J - z_J) \otimes a_J) dz_J \]

\[ = L. \]

Thus, the two forms are equal.

\[ \square \]

References

1. Aubert, G., Aujol, J.-F.: A variational approach to removing multiplicative noise. SIAM J. Appl. Math. 68, 925–946 (2008)
2. Borcea, L., Ortiz, M.: A multisattering series for impedance tomography in layered media. Inverse Probl. 15, 515 (1999)
3. Boyd, S., El Ghaoui, L., Feron, E., Balakrishnan, V.: Linear Matrix Inequalities. SIAM, Philadelphia (1994)
4. Bui-Thanh, T., Ghattas, O., Martin, J., Stadler, G.: A computational framework for infinite-dimensional Bayesian inverse problems part I: the linearized case, with application to global seismic inversion. SIAM J. Sci. Comput. 35, A2494–A2523 (2013)
5. Burckhardt, C.B.: Speckle in ultrasound b-mode scans. IEEE Trans. Son. Ultrason. 25, 1–6 (1978)
6. Calvetti, D., Dunlop, M.M., Somersalo, E., Stuart, A.M.: Iterative updating of model error for Bayesian inversion. arXiv preprint arXiv:1707.04246 (2017)
7. Calvetti, D., Ernst, O., Somersalo, E.: Dynamic updating of numerical model discrepancy using sequential sampling. Inverse Probl. 30, 110301 (2014)
8. Calvetti, D., Somersalo, E.: Bayesian image deblurring and boundary effects. In: SPIE-The International Society for Optical Engineering, vol. 5910, pp. 5910–5910–9 (2005)
9. Calvetti, D., Somersalo, E.: An Introduction to Bayesian Scientific Computing: Ten Lectures on Subjective Computing, vol. 2. Springer, New York (2007)
10. Daon, Y., Stadler, G.: Mitigating the influence of the boundary on PDE-based covariance operators. Inverse Probl. Imaging 12, 1083–1102 (2018)
11. Durand, S., Fadili, J., Nikolova, M.: Multiplicative noise removal using Landweber fidelity on frame coefficients. J. Math. Imaging Vis. 36, 201–226 (2010)
12. Fouche, S., Benie, G.B., Boucher, J.M.: Multiscale MAP filtering of SAR images. IEEE Trans. Image Process. 10, 49–60 (2001)
13. García-Lágero, M., Hermoso-Carazo, A., Linares-Perez, J.: Derivation of centralized and distributed filters using covariance information. Comput. Stat. Data Anal. 55, 312–323 (2011)
14. Guo, H., Odegard, J.E., Lang, M., Gopinath, R.A., Selesnick, I.W., Burrus, C.S.: Wavelet based speckle reduction with application to SAR based ADT/R. In: Proceedings of 1st International Conference on Image Processing, vol. 1, pp. 75–79 (1994)
15. Huang, Y., Ng, M., Zeng, T.: The convex relaxation method on deconvolution model with multiplicative noise. Commun. Comput. Phys. 13, 1066–1092 (2013)
16. Huttunen, J., Kaipio, J.: Approximation errors in nonstationary inverse problems. Inverse Probl. Imaging 1, 77–93 (2007)
17. Kaipio, J., Kolehmainen, V.: Bayesian theory and applications. In: Approximate Marginalization Over Modeling Errors and Uncertainties in Inverse Problems, pp. 644–672. Oxford University Press (2013)
18. Kaipio, J., Somersalo, E.: Statistical and Computational Inverse Problems. Springer, New York (2005)
19. Kaipio, J., Somersalo, E.: Statistical inverse problems: discretization, model reduction and inverse crimes. J. Comput. Appl. Math. 198, 493–504 (2007)
20. Kristensen, U., Scarabosio, L., Swierczynski, P., Ullmann, E., Wohlmuth, B.: Analysis of boundary effects on PDE-based sampling of Whittle–Matérn random fields. SIAM/ASA J. Uncertain. Quant. 7, 948–974 (2019)
21. Kolehmainen, V., Tarvainen, T., Arridge, S.R., Kaipio, J.P.: Marginalization of uninteresting distributed parameters in inverse problems-application to optical tomography. Int. J. Uncertain. Quant. 1, 1–17 (2011)
22. Krissian, K., Westin, C.-F., Kikinis, R., Vosburgh, K.G.: Oriented speckle reducing anisotropic diffusion. IEEE Trans. Image Process. 16, 1412–1424 (2007)
23. Lehikoinen, A., Finsterle, S., Voutilainen, A., Heikkinen, L., Vauhkonen, M., Kaipio, J.: Approximation errors and truncation of computational domains with application to geophysical tomography. Inverse Probl. Imaging 1, 371–389 (2007)
24. Lehikoinen, A., Huttunen, J., Voutilainen, A., Finsterle, S., Kowalsky, M., Kaipio, J.P.: Dynamic inversion for hydrological process monitoring with electrical resistance tomography under model uncertainties. Water Resour. Res. (in press) (2009)
25. Li, F., Ng, M.K., Shen, C.: Multiplicative noise removal with spatially varying regularization parameters. SIAM J. Imaging Sci. 3, 1–20 (2010)
26. Lipponen, A., Seppnen, A., Kaipio, J.: Reduced-order estimation. Comput. Stat. Data Anal. 312–323 (2011)
27. Michailovich, O.V., Tannenbaum, A.: Despeckling of medical ultrasound images. IEEE Trans. Ultrason. Ferroelectr. Freq. Control 53, 64–78 (2006)
28. Nicholson, R., Petra, N., Kaipio, J.P.: Estimation of the Robin coefficient field in a Poisson problem with uncertain conductivity field. Inverse Probl. 34, 115005 (2018)
29. Nissinen, A., Kolehmainen, V., Kaipio, J.P.: Compensation of modelling errors due to unknown domain boundary in electrical tomography. Inverse Probl. Imaging 10, 074010 (2016)
30. Nissinen, A., Kolehmainen, V.: Approximation errors due to unknown domain boundary in electrical tomography. Inverse Probl. Imaging 10, 058301 (2016)
31. Nissinen, A., Kolehmainen, V., Kaipio, J.P.: Marginalization of uninteresting distributed parameters in inverse problems-application to optical tomography. Int. J. Uncertain. Quant. 1, 1–17 (2011)
32. Nissinen, A., Kolehmainen, V., Kaipio, J.P.: Marginalization of uninteresting distributed parameters in inverse problems-application to optical tomography. Int. J. Uncertain. Quant. 1, 1–17 (2011)
33. Nissinen, A., Kolehmainen, V.: Approximation errors due to unknown domain boundary in electrical tomography. Inverse Probl. Imaging 10, 058301 (2016)
34. Nissinen, A., Kolehmainen, V., Kaipio, J.P.: Marginalization of uninteresting distributed parameters in inverse problems-application to optical tomography. Int. J. Uncertain. Quant. 1, 1–17 (2011)
impedance tomography. IEEE Trans. Med. Imaging 30, 231–242 (2011)
30. Rodríguez, P.: Total variation regularization algorithms for images corrupted with different noise models: a review. J. Electr. Comput. Eng. 2013, 10:10–10:10 (2013)
31. Rohatgi, V.K., Saleh, A.K.M.E.: Multiple Random Variables, pp. 99–171. Wiley, New York (2015)
32. Roininen, L., Huttunen, J.M., Lasanen, S.: Whittle–Matérn priors for Bayesian statistical inversion with applications in electrical impedance tomography. Inverse Probl. Imaging 8, 561–586 (2014)
33. Rue, H., Held, L.: Gaussian Markov Random Fields: Theory and Applications. Chapman & Hall/CRC Monographs on Statistics & Applied Probability. CRC Press, Boca Raton (2005)
34. Shi, J., Osher, S.: A nonlinear inverse scale space method for a convex multiplicative noise model. SIAM J. Imaging Sci. 1, 294–321 (2008)
35. Steidl, G., Teuber, T.: Removing multiplicative noise by Douglas–Rachford splitting methods. J. Math. Imaging Vis. 36, 168–184 (2010)
36. Stuart, A.M.: Inverse problems: a Bayesian perspective. Acta Numer. 19, 451–559 (2010)
37. Tarantola, A.: Inverse Problem Theory and Methods for Model Parameter Estimation. Society for Industrial and Applied Mathematics, Philadelphia (2004)
38. Tarvainen, T., Kolehmainen, V., Pulkkinen, A., Vauhkonen, M., Schweiger, M., Arridge, S.R., Kaipio, J.P.: Approximation error approach for compensating modelling errors between the radiative transfer equation and the diffusion approximation in diffuse optical tomography. Inverse Probl. 26, 015005 (2010)
39. Teuber, T., Lang, A.: A new similarity measure for nonlocal filtering in the presence of multiplicative noise. Comput. Stat. Data Anal. 56, 3821–3842 (2012)
40. Tison, C., Nicolas, J.M., Tupin, F., Maître, H.: A new statistical model for Markovian classification of urban areas in high-resolution SAR images. IEEE Trans. Geosci. Remote Sens. 42, 2046–2057 (2004)
41. Wong, A., Mishra, A., Bizheva, K., Clauss, D.A.: General Bayesian estimation for speckle noise reduction in optical coherence tomography retinal imagery. Opt. Express 18, 8338–8352 (2010)
42. Yin, D., Gu, Y., Xue, P.: Speckle-constrained variational methods for image restoration in optical coherence tomography. J. Opt. Soc. Am. A 30, 878–885 (2013)
43. Zhang, X., Chatwin, C., Barber, D.C.: A feasibility study of a rotary planar electrode array for electrical impedance mammography using a digital breast phantom. Physiol. Meas. 36, 1311–1335 (2015)
44. Zhao, X.-L., Wang, F., Ng, M.K.: A new convex optimization model for multiplicative noise and blur removal. SIAM J. Imaging Sci. 7, 456–475 (2014)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

R. Nicholson received his Ph.D. degree from the University of Auckland, Auckland, New Zealand, after which he spent some time in the USA. He then returned to the University of Auckland where he is currently a lecturer in the Engineering Science department. His mathematical interests include large-scale inverse problems, uncertainty quantification, modelling, numerical methods for PDEs, and linear algebra. In his spare time, he enjoys spending time with family and friends and playing all sorts of outdoor sports.

J. P. Kaipio received the M.Sc. and Ph.D. degrees from the University of Kuopio, Kuopio, Finland. He is currently a Professor of applied mathematics with the Department of Mathematics, The University of Auckland, Auckland, New Zealand. He also leads the Inverse Problems Research Group, University of Eastern Finland, Kuopio, as a part-time Professor of computational physics. He has coauthored about 140 papers, mostly in inverse problems, as well as the book Statistical and Computational Inverse Problems with Erkki Somersalo. Dr. Kaipio is a Fellow of the Institute of Physics, UK.