

# A framework for petrophysically and geologically guided geophysical inversion using a dynamic Gaussian mixture model prior

Thibaut Astic<sup>ID</sup> and Douglas W. Oldenburg

*Geophysical Inversion Facility, University of British Columbia, Vancouver, BC, Canada. E-mail:* [tastic@eoas.ubc.ca](mailto:tastic@eoas.ubc.ca)

Accepted 2019 August 24. Received 2019 August 14; in original form 2019 February 27

## SUMMARY

We propose a new framework for incorporating petrophysical and geological information into voxel-based geophysical inversion. By developing the geophysical inverse problem from a probabilistic perspective, we redesign the objective function and the iteration steps as a suite of cyclic optimization problems in which three separate MAP optimization problems are solved using geophysical, petrophysical and geological data, respectively. By quantitatively linking these data into a single framework, we recover a final inverted model that reproduces the observed, or desired, petrophysical and geological features while fitting the geophysical data. To achieve our goal we replace the Gaussian prior, used in the Tikhonov inversion approach, by a Gaussian mixture model. After each geophysical model update, the mixture parameters (means, variances and proportions) are determined by the geophysical model and the expected characteristics of the lithologies through another optimization process using the expectation–maximization algorithm. We then classify the model cells into rock units according to the petrophysical and geological information. These two additional steps over the petrophysical and geological data result in a dynamic update of the reference model and associated weights and guide the inversion towards reproducing the expected petrophysical and geological characteristics. The resulting geophysical objective function does not require extra terms to include the additional petrophysical and geological information; this is an important distinction between our work and previous frameworks that carry out joint geophysical and petrophysical data inversion. We highlight different capabilities of our methodology by inverting magnetotelluric and direct-current resistivity data in 1-D and 2-D, respectively. Finally, we apply our framework to inverting airborne frequency domain data, acquired in Australia, for the detection and characterization of saline contamination of freshwater.

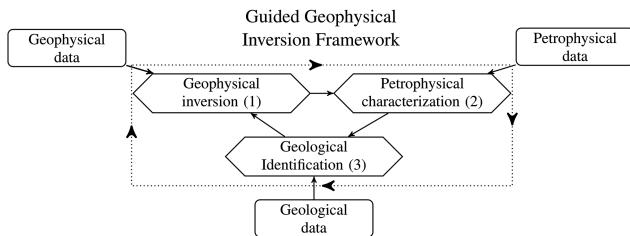
**Key words:** Electrical properties; Non-linear electromagnetics; Inverse theory; Joint Inversion; Probability distributions.

## 1 INTRODUCTION

Geophysical inversions are an essential tool for mapping the subsurface. However, the image of the underground retrieved from an inversion rarely benefits from the full range of knowledge because petrophysical relationships or expected geological features are not easily incorporated into traditional geophysical inversions. Since the early 2000s, there has been an increasing interest in including this information to obtain more realistic geological interpretations (Linde *et al.* 2015; Moorkamp *et al.* 2016). Various strategies have been implemented such as adding bound constraints and preferential trends (Li & Oldenburg 2000; Williams 2008; Lelièvre *et al.* 2009; Astic & Chouteau 2019), parametrization of the expected shapes of the geological bodies (Fullagar *et al.* 2008; McMillan *et al.* 2015), inclusion of more complex non-geophysical data such as structural data (Wu 2017), reproducing petrophysical data (Bosch *et al.* 2009;

Grana & Della Rossa 2010; Sun & Li 2015; Zhdanov & Lin 2017) and training on geological images (Lochbühler *et al.* 2015).

For the inclusion of petrophysical data, the most recent frameworks have focused on using clustering techniques such as the fuzzy C-means algorithm. This was first used in Paasche & Tronicke (2007) and later expanded in Lelièvre *et al.* (2012). This approach adds a clustering term to the objective function and allows a less strict relationship between the physical properties of interest. Further expansion of the method has been carried out by Sun & Li (2015, 2016, 2017). In addition to the fuzzy C-means clustering term, they added an iterative update to the cluster centres; this starts to introduce the notion of uncertainty for the petrophysical data. For linear problems, Grana *et al.* (2017) proposed a Bayesian formulation using a fixed Gaussian mixture model as prior, which allows them to estimate the posterior distribution and then to sample from that distribution. Giraud *et al.* (2017) on their side focused on



**Figure 1.** A graphical representation for our framework. Each diamond is a MAP estimation which requires data (shown in rectangular boxes) as well as input from the other MAP estimates.

reducing geological uncertainties by linking geophysical inversion and stochastic geological modelling through petrophysical information using a fixed Gaussian mixture model to model the petrophysical information. They modified the usual least-squares objective function of the geophysical inverse problem by adding a probabilistic term, computed from the prior petrophysical and geological distributions. To do this they required fixed, and thus very strong, priors for both the petrophysical and geological information. That information however is not often well-known and may exist only at a qualitative level.

This paper presents a new framework for petrophysically and geologically guided inversion (PGI) that generalizes concepts presented in Grana *et al.* (2017), Giraud *et al.* (2017) and Sun & Li (2015). We use a Gaussian mixture model (GMM) that represents our petrophysical and geological knowledge to regularize the geophysical inversion; this is analogous to the approach used by Grana *et al.* (2017). The idea of learning the physical properties mean values described in Sun & Li (2015) is formalized and extended to the variances and proportions of the GMM. At the same time we are able to include the geological information in the GMM in a similar way as did Giraud *et al.* (2017).

Our algorithm involves three optimization problems (Fig. 1, diamond-shape nodes) over the geophysical, petrophysical and geological knowledge, that are solved cyclically. Each optimization is cast within a Bayesian formulation as a MAP (maximum *a posteriori*) estimation over a posterior distribution. The framework is iterative and the geophysical model, the petrophysical characterization and the geological identification are updated through successive cycles. We show how petrophysical and geological information can be integrated into the smallness term of a conventional regularization operator in the geophysical inversion. Consequently we can achieve our objectives by carrying out a conventional deterministic geophysical inversion without adding extra terms in the regularization term. This is an important simplification since it eliminates the need for extra weighting parameters; it also allows previously developed Tikhonov-style inversion codes to be used. Our framework brings the petrophysical information to the same level as the geophysical information and allows us to generalize the concepts of uncertainties for the petrophysical data. This also allows us to define a formal target misfit for the petrophysical and geological information that is similar to the geophysical data target misfit; this concept has been missing from previous frameworks.

The paper proceeds in the following way. We first introduce the key concepts and vocabulary in Section 2. We view the well-known Tikhonov inversion through a probabilistic lens to relate it to a MAP estimate of a posterior probability density distribution. This defines the geophysical inversion process displayed in our framework in Fig. 1 (process 1). The GMM is then introduced as a way to represent geological and petrophysical information. Section 3 focuses on

the mathematical description of the framework. We show how geological and petrophysical information can be linked in a voxel-based geophysical inversion through the smallness term by replacing the Gaussian prior in the Tikhonov approach with a Gaussian mixture model representing both the petrophysical and geological knowledge. The smallness term can thus be interpreted as a misfit for the petrophysical and geological data. This allows us to define a natural metric for determining an acceptable misfit. We also define the geological identification process (Fig. 1, process 3). Finally, to complete the framework definition, we present our approach to the dynamic petrophysical characterization (Fig. 1, process 2). Numerical solutions such as weighting strategy, convergence and pseudocode are presented in the Section 4 which is dedicated to numerical implementation. In Section 5, synthetic and field examples are used to illustrate some essential aspects of our algorithm. A 1-D magnetotelluric (MT) example with both sharp and smooth features is used to illustrate the gain made in the recovery of a geophysical model when petrophysical information is available. We also use this example to step through the various stages of the algorithm. A 2-D direct-current (DC) resistivity example is used to demonstrate our framework when petrophysical information is minimal; for instance when only an expected number of distinct units is expected. This example is also used to highlight the gain made by the inclusion of prior geological information. Finally, we highlight how this framework can be used to incorporate constraints regarding the number of geological units and narrow down the domain of possible geophysical models. We use an airborne frequency domain electromagnetics (FDEM) field example, with data acquired in the Bookpurnong area in Australia, to invoke a geological assumption about the number of units needed to characterize saltwater contamination. We conclude with a discussion about the key components of the PGI framework.

## 2 KEY CONCEPTS

In this section, we first introduce the key concepts and vocabulary specific to a probabilistic formulation of the geophysical inverse problem. We will show that the Tikhonov objective function can be expressed in a Bayesian formulation, under the assumption of Gaussian priors, as a posterior probability density function. We then present our modelization choice, a Gaussian mixture model, for representing the petrophysical information.

### 2.1 Tikhonov inversion and its probabilistic expression

The Tikhonov inversion (Tikhonov & Arsenin 1977) casts the inverse problem as an optimization problem in which an objective function  $\Phi$ , such as shown in eq. (1), is minimized. Using the same notation convention as Oldenburg & Li (2005), the goal is to find a solution  $\mathbf{m}$  that minimizes:

$$\underset{\mathbf{m}}{\text{minimize}} \Phi(\mathbf{m}) = \Phi_d(\mathbf{m}) + \beta \Phi_m(\mathbf{m}) \quad (1)$$

such that  $\Phi_d(\mathbf{m}) \leq \Phi_d^*$

In eq. (1), the vector  $\mathbf{m}$  is our geophysical model, which represents physical properties on a mesh. The term  $\Phi_d$  is the data misfit,  $\Phi_m$  is the model regularization function and  $\beta$  is a positive scalar that adjusts the relative weighting between the two terms. A value of  $\beta$  is sought so that the data misfit  $\Phi_d$  is below an acceptable target misfit  $\Phi_d^*$  (Parker 1977).

The geophysical data misfit  $\Phi_d(\mathbf{m})$  is defined using a least-squares norm:

$$\Phi_d(\mathbf{m}) = \frac{1}{2} \|W_d(\mathcal{F}[\mathbf{m}] - \mathbf{d}_{\text{obs}})\|_2^2, \quad (2)$$

where  $\mathcal{F}$  is the forward modelling operator and  $\mathbf{d}_{\text{obs}}$  are the observed data. The matrix  $W_d$  contains the information about the data uncertainties but it is usually assumed that the data errors are uncorrelated and Gaussian.  $W_d$  then becomes a diagonal matrix with elements  $\epsilon_p^{-1}$  where  $\epsilon_p$  is the standard deviation for the  $p$ th datum.

The regularization term contends with the non-uniqueness of the inverse problem. Here we decompose it as:

$$\Phi_m(\mathbf{m}) = \alpha_s \Phi_s(\mathbf{m}) + \sum_{i \in \{x, y, z\}} \alpha_i \Phi_i(\mathbf{m}), \quad (3)$$

where  $\Phi_s$  is the smallness term that enforces similarity between the model  $\mathbf{m}$  and a reference model  $\mathbf{m}_{\text{ref}}$ . The smoothness terms  $\{\Phi_i\}$  are designed to penalize roughness of the geophysical model along the  $\{x, y, z\}$  directions. The  $\{\alpha\}$  scalar parameters are weighting parameters for the different parts of the regularization function. An extensive interpretation and usage of the various parameters in this form of regularization can be found in Lelièvre *et al.* (2009) or in Williams (2008).

Both smallness and smoothness are important in our work but the smallness term will play a dynamic role. Our goal is to constrain the value that can be possibly taken by each cell of the model. The smallness term is thus the one of particular interest here (whereas the smoothness terms constrain the transition from one cell to another). We write the smallness term as:

$$\Phi_s(\mathbf{m}) = \frac{1}{2} \|W_s(\mathbf{m} - \mathbf{m}_{\text{ref}})\|_2^2, \quad (4)$$

where the matrix  $W_s$  expresses a certain level of confidence in the reference model locally. At locations in the model domain where  $W_s$  is large then differences between the recovered model and the reference model are highly penalized and vice versa.

The same objective function as in eq. (1) can equivalently be expressed in a Bayesian formulation as a posterior probability density distribution  $\mathcal{P}(\mathbf{m}|\mathbf{d}_{\text{obs}})$  (Tarantola 2005). A posterior distribution, such as defined by the Bayes rule, is proportional to a likelihood distribution  $\mathcal{P}(\mathbf{d}_{\text{obs}}|\mathbf{m})$ , representing the data misfit, times a prior distribution  $\mathcal{P}(\mathbf{m})$  (eq. 5), normalized by the multiplicative constant  $1/\mathcal{P}(\mathbf{d}_{\text{obs}})$ :

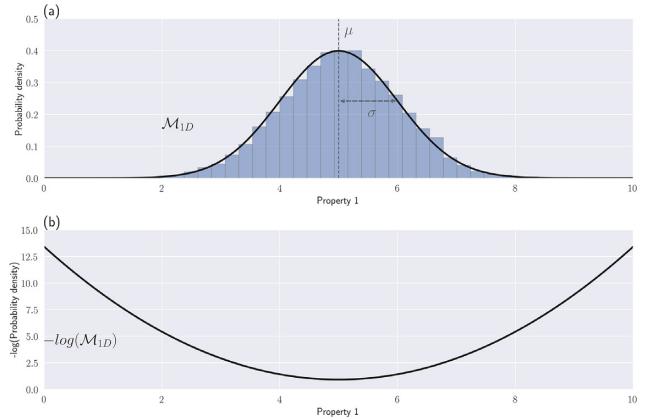
$$\mathcal{P}(\mathbf{m}|\mathbf{d}_{\text{obs}}) = \frac{\mathcal{P}(\mathbf{d}_{\text{obs}}|\mathbf{m})\mathcal{P}(\mathbf{m})}{\mathcal{P}(\mathbf{d}_{\text{obs}})} \propto \mathcal{P}(\mathbf{d}_{\text{obs}}|\mathbf{m})\mathcal{P}(\mathbf{m}) \quad (5)$$

with:

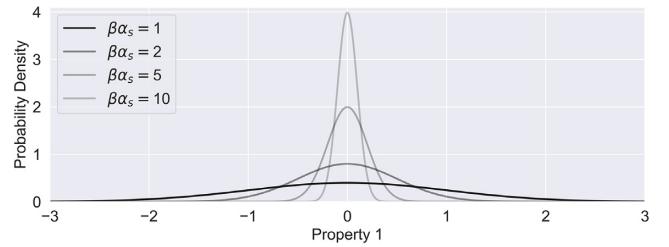
$$\mathcal{P}(\mathbf{m}) = \mathcal{P}_s(\mathbf{m})\mathcal{P}_{\{x, y, z\}}(\mathbf{m}). \quad (6)$$

The prior distribution represents the knowledge on the geophysical model such as a reference model, a certain level of smoothness, etc. (eq. 6). The model  $\mathbf{m}$  that maximizes this posterior distribution (eq. 5) is called a MAP estimate; it is the same model that minimizes the objective function. In contrast, a model that only maximizes a likelihood distribution, a data misfit term with no prior, is usually referred to as a maximum likelihood estimate (MLE).

The objective function formulation is obtained by taking the negative natural logarithm of the posterior distribution. In a Tikhonov inversion, the data misfit as well as the priors are generally expressed as least-squares (eqs 2 and 4). This translates into multivariate Gaussian distributions for the terms of the posterior distribution (see Fig. 2a for an unidimensional example). For a vector of parameters, the multivariate Gaussian distribution, denoted by  $\mathcal{N}$ , with mean  $\mu$  and covariance  $\Sigma$  (denoted  $\sigma^2$  for a unidimensional Gaussian) is



**Figure 2.** (a) Examples of a Gaussian distribution in 1-D for one physical property and (b) its quadratic negative-log equivalent.



**Figure 3.** Gaussian distributions for various confidence parameters.

defined by:

$$\mathcal{N}(\mathbf{m}|\mu, \Sigma) = \frac{\exp(-\frac{1}{2}\|\Sigma^{-\frac{1}{2}}(\mathbf{m} - \mu)\|_2^2)}{\sqrt{(2\pi)^n \det(\Sigma)}} \quad (7)$$

From eq. (7), we can see that choosing Gaussian distributions as priors and likelihood leads to a least-squares form of the terms in eq. (1), through the application of a negative natural logarithm, as displayed in eqs (2) or (4). This transformation is illustrated in Fig. 2. The summation of those terms shown in eq. (1) is simply a consequence of the fundamental property of the logarithm function, the multiplications become additions. A detailed derivation can be found in Appendix A.

In this context, the prior distribution and the regularization are equivalent. The prior distribution  $\mathcal{P}_s$  is equivalent to the smallness term defined in 4. It can be expressed as a multivariate Gaussian distribution. Its mean is the reference model  $\mathbf{m}_{\text{ref}}$  and its covariance is the weighting matrix  $(W_s^T W_s)^{-1}$ . The factor  $\beta\alpha_s$  expresses our confidence in this prior (eq. 8). Increasing this confidence parameter  $\beta\alpha_s$  is equivalent to reducing the variance around the reference model and thus the recovered solution becomes closer to the reference (such as illustrated in Fig. 3).

$$\mathcal{P}_s(\mathbf{m}) = \mathcal{N}(\mathbf{m}|\mathbf{m}_{\text{ref}}, (\beta\alpha_s W_s^T W_s)^{-1}) \quad (8)$$

or equivalently:

$$\mathcal{P}_s(\mathbf{m}) \propto \mathcal{N}(\mathbf{m}|\mathbf{m}_{\text{ref}}, (W_s^T W_s)^{-1})^{\beta\alpha_s} \quad (9)$$

Defining the smallness prior as a least-squares term such as in eq. (4) thus assumes a Gaussian distribution of the model  $\mathbf{m}$  values around the reference model. This is unlikely to correspond to the true physical property distribution, especially when the initial model is a half-space as commonly used. With the Tikhonov approach, the incorporation of knowledge about the physical properties can only

be done locally at fixed locations defined by the users (Lelièvre *et al.* 2009). A more versatile method is required.

Our focus is on designing a more meaningful and dynamically updatable prior for the smallness term  $\mathcal{P}_s(\mathbf{m})$  when non-geophysical data are available but not necessarily well located. This will allow us to guide the model  $\mathbf{m}$  towards producing an expected petrophysical distribution as well as capturing geological features. The inclusion of various types of information is done through the prior distribution. Our goal is thus to characterize each optimization process in Fig. 1 as a MAP estimate. We have addressed the geophysical inversion problem (Fig. 1, process 1). We now focus on the petrophysical characterization and geological identification.

## 2.2 Gaussian mixture model representation of petrophysical information

To include physical property information into the inversion, we first need to define a way to model the physical property distributions. Our choice is to use a Gaussian mixture model, where each geological unit's physical property is represented by a Gaussian distribution.

Consider that we are given a petrophysical data set. We have  $n$  samples, denoted  $\{s_i, i = 1..n\}$ , which represent the measured physical properties. For each sample we know its current geological classification. This is denoted by  $j = \{1..c\}$  where  $c$  is the number of distinct geological units. We can use the geological classification as a categorical variable. A sample  $i$  belonging to unit  $j$  is noted  $s_{i \in j}$ .

For each geological unit  $j$ , we can fit a Gaussian probability density distribution. If the unit does not follow a Gaussian distribution, we can often project the physical properties into a transformed space where it appears approximately Gaussian, such as a log-space for electrical conductivity or magnetic susceptibility [see the notion of mapping in the SimPEG framework as defined in Cockett *et al.* (2015) and Kang *et al.* (2015)]. With known labels, the MLE of the Gaussian distribution parameters for each unit  $j = \{1..c\}$ , its mean  $\mu_j$  and its variance  $\sigma_j^2$  (or standard deviation  $\sigma_j$ ), plus its proportion  $\pi_j$  among the data set, are given in eqs (10)–(12):

$$\pi_j = \frac{n_j}{n} \quad (10)$$

$$\mu_j = \frac{1}{n_j} \sum_{i=1}^{n_j} s_{i \in j} \quad (11)$$

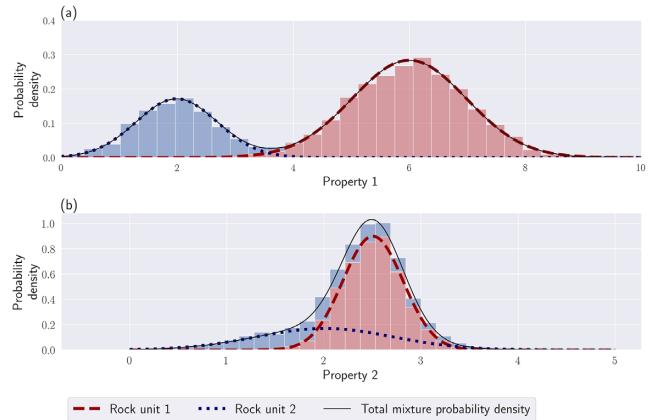
$$\sigma_j^2 = \frac{1}{n_j} \sum_{i=1}^{n_j} (s_{i \in j} - \mu_j)^2, \quad (12)$$

where the  $s_{i \in j}$  are our  $n_j$  physical property measurements for each geological unit  $j$ .

The full probability function to observe an unlabeled physical property data point  $s_i$  can be written as a GMM (see also Figs 4a and b):

$$\mathcal{P}(s_i | \Theta) = \sum_{j=1}^c \pi_j \mathcal{N}(s_i | \mu_j, \sigma_j^2) \quad (13)$$

The variable  $\Theta$  holds the GMM global variables  $\Theta = \{\pi_j, \mu_j, \sigma_j^2\}_{j=1..c}$ . A GMM is a parametric probability density distribution that can fit any continuous probability density distribution, when its number of clusters tends to infinity (Murphy 2012).



**Figure 4.** Examples of a Gaussian mixture like samples sets. (a) A GMM with two distinct units (the approximation in eq. (20) is valid). (b) A GMM with overlapping units, the approximation in eq. (20) is not valid.

It has gained popularity in recent years for representing geological and petrophysical information (Grana & Della Rossa 2010; Granek 2011; Giraud *et al.* 2017; Grana *et al.* 2017).

We denote  $\mathbf{z}$  as our categorical variable for the labels, i.e. our geological classification. We call this variable our membership. Given an unlabeled data point  $s_0$ , its membership  $z_0$  takes the value of the geological unit with the highest probability (see eq. 14).

$$z_0 = \underset{z \in \{1..c\}}{\operatorname{argmax}} \mathcal{P}(s_0 | z) \mathcal{P}(z) \quad (14)$$

with:

$$\mathcal{P}(z) = \pi_z \text{ and} \quad (15)$$

$$\mathcal{P}(s_0 | z) = \mathcal{N}(s_0 | \mu_z, \sigma_z^2). \quad (16)$$

Note that eq. (14) defines a MAP estimate of the geological identifier  $\mathbf{z}$ . The parameters  $\{\pi_j\}_{j=1..c}$  represent our prior expectation of observing a certain unit before observing  $s_0$  (eq. 15). This MAP estimate will be our template for designing the geological identification process in our framework (Fig. 1, process 3).

The only posterior distribution left to define is for the petrophysical characterization process. We wish to determine MAP estimates of the GMM parameters when labels are unknown (Fig. 1, process 2). This is done in Section 3.2.

## 3 GMM AS SMALLNESS PRIOR FOR GEOPHYSICAL INVERSION

### 3.1 GMM prior definition

Suppose we are given petrophysical and geological information about the geophysical model that we want to recover. Petrophysical information can include mean or variance values of physical properties for different geological units. Geological information can consist of an expected number of distinct units or an anticipation of encountering certain rock units in particular locations. We need to design a prior that offers the maximum flexibility for representing that petrophysical and geological prior knowledge.

In eq. (17) we propose a GMM that is designed to serve as a prior on  $\mathbf{m}$ , whose parameters are both spatially (index  $i$ ) and

lithologically (index  $j$ ) dependent. Our probability function is:

$$\mathcal{M}(\mathbf{m}|\Theta) = \prod_{i=1}^n \sum_{j=1}^c \mathcal{P}(z_i = j) \mathcal{N}(m_i|\mu_j, w_i^{-2}\sigma_j^2) \quad (17)$$

where:

- (i)  $c$  is the number of distinct rock units, or clusters.
- (ii)  $n$  is the number of active cells in the mesh.
- (iii)  $m_i$  represents our physical property value at the  $i^{\text{th}}$  cell.
- (iv)  $\mathbf{z}$  describes the membership to a certain rock unit.
- (v)  $\mathcal{P}(z_i = j)$  represents our geological information. It is the prior probability of observing rock unit  $j$  at location  $i$  (vector notation:  $\mathcal{P}(\mathbf{z})$ ). It can be either constant over the whole area, then representing an expected relative volume of each unit denoted by the global proportion  $\pi_j$ , or locally set if the information is available [e.g. borehole logs, outcrops or geological modelling (see Giraud *et al.* 2017)].
- (vi)  $\mu_j$  is the mean physical property of rock unit  $j$ .
- (vii)  $w_i^{-2}\sigma_j^2$  is the variance of the rock unit  $j$  at the  $i^{\text{th}}$  cell. It includes both the expected physical property variance  $\sigma_j^2$  of the rock unit  $j$  globally and a local prior confidence  $w_i^2$  that, at a certain cell  $i$ , the model value  $m_i$  should belong closer or farther away from the mean  $\mu_j$  of the rock unit  $j$ . Those  $\mathbf{w}$  weights can be used to include depth or sensitivity weighting.
- (viii)  $\Theta$  holds the GMM global variables  $\Theta = \{\pi_j, \mu_j, \sigma_j^2\}_{j=1..c}$ .

Instead of adding an extra term in the objective function, we use the previously defined GMM probability distribution (eq. 17), representing our current geological and petrophysical knowledge, as our smallness prior. To have consistency with the Tikhonov framework, such as that defined in eq. (8), we write:

$$\mathcal{P}_s(\mathbf{m}) \propto \mathcal{M}(\mathbf{m}|\Theta)^{\beta\alpha_s}. \quad (18)$$

To formulate our objective function, we apply a negative logarithm to the new posterior distribution, which uses the GMM probability distribution as the smallness prior (eq. 18). The resulting smallness term is then:

$$\Phi_s(\mathbf{m}) = - \sum_{i=1}^n \log \left( \sum_{j=1}^c \mathcal{P}(z_i = j) \mathcal{N}(m_i|\mu_j, w_i^{-2}\sigma_j^2) \right). \quad (19)$$

Although we have had positive results with using the above *LogSumExp* function in the smallness regularizer (the exponential function being in the Gaussian distribution  $\mathcal{N}$ ), there are benefits of using the following approximation. Gaussian distributions decay exponentially. If the units have distinguishable enough contrasts, it is reasonable to approximate the *LogSumExp* function in eq. (19) by its locally dominant quadratic term. This means considering only the most probable rock unit for each cell given the current model  $\mathbf{m}$ . The other terms of the summation are deemed negligible. For example consider the 1D physical property distributions in Fig. 4. For the physical property distribution in Fig. 4(a), the individual Gaussian functions representing each rock unit can locally approximate the GMM extremely well. On the other hand, for the physical property distribution displayed in Fig. 4(b), there is too much overlap between the two individual Gaussian components. Individually they are not a good approximation to the full GMM distribution and the above assumption is incorrect. In that case, one can either work with the full expression shown in eq. (19), or potentially re-evaluate whether the individual distributions are distinct enough to be considered as separate units. As a rule of thumb, contrasts are deemed distinguishable enough to apply this approximation if the centres of

two clusters are separated by at least three times the largest standard deviation. By using the parameters of the most likely cluster at each active cell we can approximate  $\Phi_s(\mathbf{m})$  as:

$$\Phi_s(\mathbf{m}) = \frac{1}{2} \|W_s(\Theta, \mathbf{z}^*)(\mathbf{m} - \mathbf{m}_{\text{ref}}(\Theta, \mathbf{z}^*))\|_2^2 \quad (20)$$

with:

$$\mathbf{z}^* = \underset{\mathbf{z}}{\operatorname{argmax}} \mathcal{P}(\mathbf{m}|\mathbf{z}) \mathcal{P}(\mathbf{z}) \quad (21)$$

$$\mathbf{m}_{\text{ref}}(\Theta, \mathbf{z}^*) = \mu_{\mathbf{z}^*} \quad (22)$$

$$W_s(\Theta, \mathbf{z}^*) = \text{diag}(\mathbf{w} \circ \sigma_{\mathbf{z}^*}^{-1}). \quad (23)$$

Note the similarity of the smallness term in eq. (20) with the Tikhonov formulation (eq. 4). This allows us to generate a petrophysically and geologically guided inversion that uses traditional algorithms, and thus can rely on the research literature on the Tikhonov approach.

Our approach, regarding the regularized geophysical problem, can be summarized as follows: At each iteration, we first identify the most probable unit for each cell; we call it its membership, and store it in the categorical vector  $\mathbf{z}$  (eq. 21). The reference model  $\mathbf{m}_{\text{ref}}$  and smallness weights  $W_s$  are then updated according to the membership  $\mathbf{z}$ . The reference model at each cell takes the physical property mean value of the most likely unit  $z_i^*$  (eq. 22). The smallness weights depend on the value of the corresponding variances and on our local weights (e.g. depth or sensitivity weighting; see eq. (23)). These updates represent the newly acquired knowledge on  $\mathbf{m}$  according to the petrophysical distribution  $\Theta$  and the geological information  $\mathcal{P}(\mathbf{z})$ . We refer to the updated reference model as the learned reference model. The next inversion step then pushes each cell of the mesh towards the mean of its most probable rock unit. The strength of the push is proportional to the variance of the physical property for that unit (the higher the variance, the less we push), and our weights  $\mathbf{w}$ .

Eq. (21) defines the MAP process for the geological identification (Fig. 1, process 3). Eqs (22) and (23) link the petrophysical and geological information with the geophysical inversion.

One could proceed by using the means and covariances fitted on the petrophysical data, as well as the proportions from the geological data, to compute  $\mathbf{z}$ ,  $\mathbf{m}_{\text{ref}}$  and  $W_s$  at each iteration (eqs 21, 22 and 23). However when this knowledge is incomplete or uncertain, our estimated GMM may not be accurate. Carter-McAuslan *et al.* (2015) have investigated that issue and shown that inaccurate petrophysical information can significantly affect the recovered model. Sun & Li (2015, 2016, 2017) partially overcome this by updating the cluster centres through the iteration process; they average the observed means in the geophysical model with the means from the petrophysical measurements through a separate clustering optimization problem. We generalize this approach to all of the parameters of the GMM. In the next Section 3.2, we formulate the petrophysical characterization step of our framework (Fig. 1, process 2) as a MAP estimation process.

### 3.2 Updating the GMM

One can either consider that the GMM parameters derived from petrophysical and geological data give the best estimation of the true distribution, or that it is only an approximation, and additional information can be obtained from the geophysical model during the

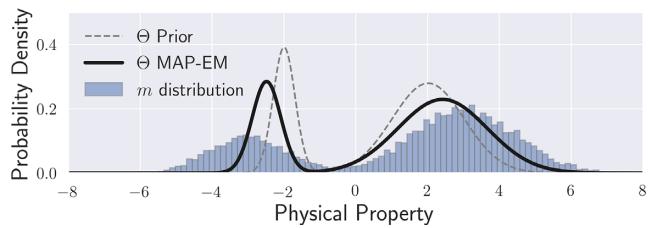
inversion (Sun & Li 2015). We prefer this later approach. Since the prior knowledge might be incomplete or only known qualitatively, it is important to mitigate possible biases that can affect the final recovered model (Carter-McAuslan *et al.* 2015). For that purpose Sun & Li (2015) updated each cluster centre of their fuzzy C-means clusters through a weighted average of the expected centre from the petrophysical data with the observed centre in the geophysical model. From a statistical view point, this resembles the use of a MAP estimator to find the cluster means applied to the geophysical model itself, with the addition of a conjugate prior based on the petrophysical data. The notion of conjugate prior was first introduced in Raiffa & Schlaifer (1961). By definition, a prior is called a conjugate prior to the likelihood function if the posterior belongs to the same family of probability distributions as the prior. For example, the Gaussian distribution is self-conjugate. This means that if the likelihood follows a Gaussian distribution, choosing a Gaussian prior ensures that the posterior also belongs to the same family of probability density distributions. We use the conjugate prior approach to formalize the update of the means done in Sun & Li (2015) and generalize it to the other GMM parameters.

Obtaining the GMM parameters  $\Theta$  is thus made part of the iterative inversion process. We define the petrophysical characterization process (Fig. 1, step 2) as a MAP estimator for  $\Theta$ . This MAP estimate benefits both from the petrophysical and geological data known prior to the inversion and the current geophysical model. After each geophysical inversion iteration, and before updating the membership  $\mathbf{z}$ , we learn the mixture parameters  $\Theta$  using the current geophysical model  $\mathbf{m}$  and our confidence in the prior knowledge  $\Theta_{\text{prior}}$ . We define a posterior distribution on  $\Theta$ , given the geophysical model  $\mathbf{m}$  and prior knowledge, as:

$$\mathcal{P}(\Theta|\mathbf{m}) \propto \mathcal{M}(\mathbf{m}|\Theta)\mathcal{P}(\Theta). \quad (24)$$

In Section 2.2, the problem of fitting a single Gaussian distribution to each unit, and thus defining its petrophysical mean and covariance, was made easy by knowing the geological unit of each sample. A much harder problem is to recover the full multimodal probability function described in eq. (13) when there are missing labels. For example, in an inverted geophysical model  $\mathbf{m}$ , most locations for the geological units are unknown except where we have drillholes and outcrops. In fact, this is usually the information we are trying to determine. The categorical variable  $\mathbf{z}$  that represents our geological classification is now a hidden variable that we want to find along with the GMM parameters. Clustering algorithms are especially designed for this sort of task. The expectation–maximization (EM) Algorithm (Dempster *et al.* 1977) is one of the most widely used algorithms for semi-supervised and unsupervised learning in mixture modelling. It allows us to fit the parameters  $\Theta = \{\pi_j, \mu_j, \sigma_j^2\}_{j=1..c}$  of a mixture model to a multimodal distribution with partial or no labeling. The EM algorithm in its original form is a MLE, thus no prior information is involved. To include prior information in the algorithm, and thus find a MAP estimate for the posterior distribution on  $\Theta$  shown in eq. (24), we use a MAP variation of the EM algorithm (also introduced in Dempster *et al.* 1977); we refer to this algorithm as the MAP-EM algorithm.

For the MAP-EM algorithm, we need to choose how to model the prior  $\mathcal{P}(\Theta)$  for each type of parameter. We can follow either a conjugate prior or semi-conjugate prior approach as described by Murphy (2012). Each step of the MAP-EM algorithm applied to a GMM with a conjugate or semi-conjugate prior can be understood as a weighted average of the observed parameters with their priors. Fig. 5 shows an example of such a MAP estimate of the GMM



**Figure 5.** Illustration of fitting a Gaussian mixture model with parameters  $\Theta$  through the MAP-EM algorithm, given the geophysical model  $\mathbf{m}$  and the prior petrophysical distribution  $\Theta_{\text{prior}}$  with all confidences set to unity.

parameters. We define the input of each prior function so that an input value of unity denotes an equal weighting of the observed and prior distributions. We refer to those input values as our confidences in the GMM prior parameters. Confidences of zero in the prior generate a MLE estimator; the prior parameters are not taken into account and the algorithm reverts to the normal EM algorithm. Infinite confidences in the prior fix the GMM's parameters equal to their priors value.

The MAP-EM algorithm is an iterative process. We initialize by starting with the parameters determined at the previous cycle. The usual stopping criteria defined in the literature for the MAP-EM algorithm is to define a minimum increase of the posterior probability density value. At each MAP-EM iteration ( $k$ ), we first compute the responsibility  $\{n_{ij}\}_{j=1..c, i=1..n}$  of each cluster  $j$  for each point  $m_i$ . This is referred as the E-step in the EM algorithm and it stays unchanged in its MAP-EM variation:

$$n_{ij}^{(k)} = \frac{\mathcal{P}(z_i = j)^{(k-1)} \mathcal{N}(m_i | \mu_j^{(k-1)}, (\sigma_j^2)^{(k-1)})}{\sum_{t=1}^c \mathcal{P}(z_i = t)^{(k-1)} \mathcal{N}(m_i | \mu_t^{(k-1)}, (\sigma_j^2)^{(k-1)})}. \quad (25)$$

We then compute the MAP-EM steps at iteration ( $k$ ) with respect to the global proportion weights  $\{\pi_j\}_{j=1..c}$ , means  $\{\mu_j\}_{j=1..c}$  and variances  $\{\sigma_j^2\}_{j=1..c}$  of the GMM. These steps are referred to as the M-step in the EM-algorithm. The MAP-EM variation M-step includes prior information as part of the iterative update.

The global proportion weights follow a categorical distribution whose conjugate prior is the Dirichlet distribution (eq. 26):

$$\mathcal{P}(\boldsymbol{\pi}) = \text{Dir}(\zeta \pi_{\text{prior}} V - 1). \quad (26)$$

At iteration ( $k$ ), the resulting posterior estimate of  $\{\pi_j\}_{j=1..c}$ , given  $\{\pi_{j\text{prior}}\}_{j=1..c}$  and with confidence  $\{\zeta_j\}_{j=1..c}$ , is shown in eq. (27):

$$\pi_j^{(k)} = \frac{V_j^{(k)} + \zeta_j \pi_{j\text{prior}} V}{V(1 + \sum_{t=1}^c \zeta_t \pi_{t\text{prior}})} \quad (27)$$

with:

$$V_j^{(k)} = \sum_{i=1}^n v_i n_{ij}^{(k)} \quad (28)$$

$$\text{and } V = \sum_{i=1}^n v_i, \quad (29)$$

where  $v_i$  is the volume of the  $i$ th cell and  $V$  is the volume of the active mesh. The addition of these volume weights is necessary because the cells of our mesh can have different sizes. We want to use a prior information that is mesh-independent. That is why we use volumetric proportions instead of cell counts, which is the way the EM algorithm is usually implemented.

The full conjugate priors of the likelihood function in eq. (24) for the means and variances are not independent of each other and they

follow a normal-inverse-gamma distribution (Murphy 2012). However using the full conjugate prior approach for the variances can be harmful in the first few iterations. Mean values from models at early iterations may be far from their final target values and that would drive the GMM estimated by MAP-EM to have very high variances (to compensate for the difference between the observed and prior means) and thus compromise the clustering (see Appendix B, for further details).

To avoid this issue, a semi-conjugate prior approach is possible where the dependency of the means and variances is avoided by using the conditional distributions as priors. In this case the semi-conjugate prior for the means follows a Gaussian distribution (eq. 30) while the semi-conjugate prior for the variances follows an inverse-gamma distribution (eq. 31). As usually done in the statistics literature (Murphy 2012), we use an alternative parametrization of the inverse-gamma distribution called the inverse chi-squared distribution. This re-parametrization allows for a simpler definition of the parameters. The semi-conjugate priors for the means and variances, with respective confidences  $\{\kappa\}$  and  $\{\nu\}$ , are:

$$\mathcal{P}(\mu|\sigma^2) = \mathcal{N}(\mu|\mu_{prior}, (\kappa\pi_{prior}V)^{-1}) \quad (30)$$

and:

$$\mathcal{P}(\sigma^2|\mu) = \mathcal{X}^{-2}(\sigma^2|\nu\pi_{prior}V, \sigma_{prior}^2). \quad (31)$$

Note that if we put prior information only on the means, or only on the variances, then this is a full conjugate prior. The semi-conjugate prior approach arises only when we put prior information on both means and variances (meaning that  $\{\kappa\}$  and  $\{\nu\}$  are both non zero).

The update of the means is the same for both the full conjugate and semi-conjugate prior approaches. The current ( $k$ ) posterior estimate of the means  $\{\mu_j\}_{j=1..c}$  is a weighted average of the MLE estimates  $\{\bar{m}_j\}_{j=1..c}$  from eq. (11), using the geophysical model as the samples, with the prior means  $\{\mu_{j,prior}\}_{j=1..c}$  and with confidences  $\{\kappa_j\}_{j=1..c}$ . This update to the means is shown in eq. (32):

$$\mu_j^{(k)} = \frac{V_j^{(k)}\bar{m}_j^{(k)} + \kappa_j\pi_{j,prior}V\mu_{j,prior}}{V_j^{(k)} + \kappa_j\pi_{j,prior}V} \quad (32)$$

with:

$$\bar{m}_j^{(k)} = \frac{\sum_{i=1}^n v_i n_{ij}^{(k)} m_i}{V_j^{(k)}}. \quad (33)$$

For the semi-conjugate prior, the current ( $k$ ) posterior estimate of the variances  $\{\sigma_j^2\}_{j=1..c}$ , given in eq. (34), is a weighted average, with confidence  $\{\nu_j\}_{j=1..c}$ , of the MLE estimations of the variances  $\{\sigma_{\bar{m},j}^2\}_{j=1..c}$  from eq. (12) with their prior  $\{\sigma_{j,prior}^2\}_{j=1..c}$ .

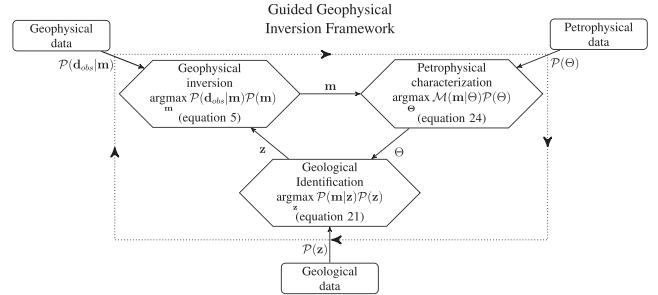
$$\sigma_j^{2(k)} = \frac{V_j^{(k)}\sigma_{\bar{m},j}^{2(k)} + \nu_j\pi_{j,prior}V\sigma_{j,prior}^2}{V_j^{(k)} + \nu_j\pi_{j,prior}V} \quad (34)$$

with:

$$\sigma_{\bar{m},j}^{2(k)} = \frac{1}{V_j^{(k)}} \sum_{i=1}^n v_i n_{ij}^{(k)} (m_i - \bar{m}_j^{(k)})^2. \quad (35)$$

We run fully the MAP-EM algorithm to learn a new petrophysical distribution with parameters  $\Theta = \{\pi_j, \mu_j, \sigma_j^2\}_{j=1..c}$  after each iteration ( $t$ ) of the geophysical inversion process, using the current geophysical model  $\mathbf{m}^{(t)}$  and the prior petrophysical distribution  $\Theta_{prior}$ .

This completes our derivation. Eq. (24) in particular describes the MAP estimator for the petrophysical characterization. Our three



**Figure 6** A graph approach illustrating how the various posterior distribution MAP estimate processes are interlocked with each other.

MAP estimates are interconnected through the sharing of the geophysical model, petrophysical distribution and geological classification. The final framework is summarized in Fig. 6, which is the completed version of the previous diagram (Fig. 1). The next step is to numerically implement this procedure. That is done in the following section.

## 4 NUMERICAL IMPLEMENTATION

Our framework defines three inverse problems that are interconnected and special care is required to ensure that each is reasonably well solved. In this section, we look at some of the essential implementation elements of this framework. A comprehensive algorithm and a flowchart can be found in Appendix C. However, before going into the details of the pseudocode, we formalize the concept of fitting the petrophysical and geological data and thus define a stopping criteria for our numerical solution.

### 4.1 Petrophysical target misfit and inversion stopping criteria

The smallness term, such as defined in eq. (20), measures how well the geophysical model fits the petrophysical and geological data through the GMM. We need to define a target for this misfit term.

For the definition and evaluation of the petrophysical target misfit, we focus on the petrophysical and geological information contained in the smallness term that is summarized in eqs (36)–(39) and identified as  $\Phi_{petro}$ .

$$\Phi_{petro}(\mathbf{m}) = \frac{1}{2} \|W_s(\Theta, \mathbf{z}^*)(\mathbf{m} - \mathbf{m}_{ref}(\Theta, \mathbf{z}^*))\|_2^2 \quad (36)$$

with:

$$\mathbf{z}^* = \operatorname{argmax}_{\mathbf{z}} \mathcal{P}(\mathbf{m}|\mathbf{z})\mathcal{P}(\mathbf{z}) \quad (37)$$

$$\mathbf{m}_{ref}(\Theta, \mathbf{z}^*) = \mu_{\mathbf{z}^*} \quad (38)$$

$$W_s(\Theta, \mathbf{z}^*) = \operatorname{diag}(\sigma_{\mathbf{z}^*}^{-1}). \quad (39)$$

At each iteration, the  $i$ th element  $(m_i - m_{ref,i})$  is a Gaussian random variable of mean zero and variance  $\sigma_i^2$ . Consequently, elements in  $W_s(\mathbf{m} - \mathbf{m}_{ref})$  are Gaussian with zero mean and unit standard deviation. Thus  $\Phi_{petro}$  is a chi-squared statistical parameter with expectation:

$$E[\Phi_{petro}] = \Phi_{petro}^* = \frac{n}{2}. \quad (40)$$

This defines the petrophysical target misfit value  $\Phi_{\text{petro}}^*$  for  $\Phi_{\text{petro}}$  as being half the number of active cells  $n$  in the geophysical model (eq. 40). As one condition for stopping our algorithm we want to find an  $\mathbf{m}$  for which  $\Phi_{\text{petro}}$  is less than or equal to  $\Phi_{\text{petro}}^*$ .

The eq. (36) for evaluating the petrophysical stopping criterion omits the weights  $\mathbf{w}$  (eq. 20). The additional weights  $\mathbf{w}$  are used to include depth or sensitivity weighting, which are more related to the physics of the survey rather than to the desired properties of the geophysical model. We have taken two approaches; the first is to minimize  $\Phi_s$  but keep the stopping criterion attached to  $\Phi_{\text{petro}}$ . The second is to minimize and measure the stopping criterion using  $\Phi_s$  after adjusting the target misfit value to be  $\|\mathbf{w}\|_2^2/2$ . Thus far we prefer the former approach because the second has lead us in certain cases to underfit the petrophysical distribution. The examples, developed in Sections 5, all use various weighting strategies and the  $\Phi_{\text{petro}}$  stopping criterion approach.

We note this stopping criterion for the petrophysical data is analogous to the geophysical data misfit criterion. The geophysical data misfit term, such as defined in eq. (2), is also a sum of Gaussian variables with zero mean and unit standard deviation. Thus a reasonable target misfit, such as introduced in Parker (1977), is:

$$\Phi_d^* = \frac{n_d}{2}, \quad (41)$$

where  $n_d$  is the number of geophysical data.

Our algorithm stops when we have found a model that simultaneously has both  $\Phi_d$  and  $\Phi_{\text{petro}}$  equal or below their respective target value  $\Phi_d^*$  and  $\Phi_{\text{petro}}^*$ . Reaching multiple target misfits adds a new challenge to solving the inverse problem. We address this in the following section as we work through the important components of the pseudocode.

## 4.2 Pseudocode and Weighting strategies

At each iteration of our PGI algorithm, three inverse problems need to be solved. Our algorithm is provided in two forms in Appendix C. The first is a pseudocode (algorithm 1), the second is a detailed flowchart that is linked to the pseudocode (Fig. C1). Here we discuss some of the important elements that assist in understanding the details of the algorithm.

### 4.2.1 Step 1: Initialization

We are solving three MAP problems. Each needs to be initialized with an initial geophysical model  $\mathbf{m}_0$ , a petrophysical distribution  $\Theta_0$  and a membership  $\mathbf{z}_0$ , respectively. We usually have these elements consistent with each other. For example, the initial reference model is usually the starting model. We can start from a background half-space  $\mathbf{m}_0$ , whose value is the same as in  $\Theta_0$ , and set  $\mathbf{z}_0$  at the background unit everywhere.

For the success of a geophysical inversion, the choice of the various weighting parameters can be critical. Our framework still uses a conventional geophysical inversion objective function, and hence previous studies and good practices are still applicable. We need to provide uncertainties  $W_d$  for the geophysical data least-squares misfit term (eq. 2). The objective function (eq. 1) requires a reference model and initial weighting parameters. A common practice is to start with a large  $\beta$  and then decrease its value during the inversion iterative process (Farquharson & Oldenburg 2004). For the smallness and derivative weights  $\{\alpha\}$ , the similarity of our formulation with the Tikhonov approach allows us to rely on

the existing literature (e.g. Williams 2008; Lelièvre *et al.* 2009; Oldenburg & Li 2005).

For the petrophysical inverse problem, we initialize the petrophysical distribution by setting  $\Theta^0 = \Theta_{\text{prior}}$ . We thus need to design  $\Theta_{\text{prior}}$ . For each unit, when petrophysical measurements or a priori estimates are available, both the mean and variance of the GMM cluster can be set. When no information is available, it is still required to provide some estimate for the variances. These prior variances can be thought of as petrophysical noise levels, analogous to the geophysical noise levels. They play an essential role in determining the petrophysical target misfit and regulate the clustering of each unit around its, to be determined, mean value. The proportions represent geological knowledge. When they are set globally, they represent the expected relative volume of each unit. The choice of confidences in the prior parameters is still a challenging question (see Section 3.2 for definitions). When the petrophysical information is well-known, values of unity, or higher, have worked well. When no petrophysical information is known *a priori*, values of unity or above for the confidences in the prior variances are used, while confidences in the means and proportions are set to zero. In the examples section (Section 5) we show examples for both the known and unknown petrophysical information cases.

The starting geological model can be determined by the starting geophysical model and petrophysical distribution. Additional local knowledge about the geology can be added through the choice of local proportion values for the GMM. We show a simple demonstration through the 2-D resistivity example (Section 5.2). For example, knowing that a unit appears at a certain location, can be added to the inversion. For that purpose we can locally set the proportion for this unit to unity, consequently giving a probability of zero for the other units at that same location.

This initialization completes step 1 in our pseudocode (algorithm 1) and flowchart (Fig. C1). The next step is to evaluate our stopping criteria and update the models.

### 4.2.2 Step 2–5: Check convergence and update the model

At step 2, we check if the convergence criteria  $\Phi_d \leq \Phi_d^*$ , and  $\Phi_{\text{petro}} \leq \Phi_{\text{petro}}^*$  have been met. If these are not satisfied, then a model update is required. We start by taking a single Gauss-Newton step on the geophysical objective function (Nocedal & Wright 2006) (step 3). Step 4 consists in running the MAP-EM algorithm described in Section 3.2 to learn a new petrophysical distribution parameters  $\Theta^{(t)}$ , using the updated geophysical model  $\mathbf{m}^{(t)}$ . Finally, at step 5 we update the membership  $\mathbf{z}^{(t)}$ , using  $\mathbf{m}^{(t)}$  and  $\Theta^{(t)}$  according to eq. (21). This updates the smallness weights and reference model for the next iteration, such as described in eqs (22) and (23).

### 4.2.3 Step 6: Updating $\beta$ and $\alpha_s$

In this step, we adjust our weights so that the inversion arrives at a solution that acceptably fits both the geophysical data and the GMM containing the geological and petrophysical data. We have two adjustable parameters,  $\beta$  and  $\alpha_s$ . The parameter  $\beta$  balances the data misfit with the total regularization function, while  $\alpha_s$  allows us to adjust the importance of the smallness term.

Our current procedure uses a computationally inexpensive heuristic approach to find suitable  $\beta$  and  $\alpha_s$  parameters such that both geophysical and petrophysical misfits are equal or below their respective target value. We initially set  $\beta$ , and  $\{\alpha\}$  to values we would

use for a standard inversion. We first invoke a  $\beta$ -cooling to get towards a solution that fits the geophysical data. Once reached, we focus more on fitting the petrophysical data by adjusting  $\alpha_s$ .

If the geophysical data misfit term reaches a plateau before attaining its target, we decrease  $\beta$  by a predetermined factor (step 6, first case scenario). We think of  $\beta$  as an outer-control parameter since it is a weighing for the entire regularization function, which includes smallness and gradient terms.

In our case we also want to find a solution that fits the petrophysical and geological data. For this we fine-tune the inner-control parameter  $\alpha_s$ . If the geophysical data misfit target  $\Phi_d^*$  has been reached, but not the petrophysical data misfit target  $\Phi_{\text{petro}}^*$ , we put more emphasis on the smallness model term. We increase  $\alpha_s$  by a factor  $\Phi_d^*/\Phi_d^{(t)}$  (step 6, second case scenario). This is similar to the  $\beta$ -warming strategy described in Fournier (2015). Generally this allows us to focus the algorithm on reaching the petrophysical data misfit target without significantly altering the geophysical data misfit. If the geophysical data misfit is increased too much, then  $\beta$  is reduced. In summary, we use  $\alpha_s$  as a fine-tuning parameter as it is not adjusted until late stages of the inversion when the major structures have already been identified.

#### 4.2.4 Step 7: Final output

The algorithm stops when both the geophysical and petrophysical misfits are below or equal to their respective targets. The PGI framework outputs three quantities: the recovered geophysical model, the learned petrophysical distribution and membership. The geophysical model fits both the geophysical data and the learned petrophysical distribution. The membership attributes a rock unit to each cell; we refer to this categorized model as a quasi-geological model, as proposed by Li *et al.* (2019).

#### 4.2.5 Options

We have the possibility to include, or not, the learned reference model inside the smoothness components of the objective function. Our preferred strategy is to leave it out until we have reached the geophysical data misfit target and when only a small fraction of the membership  $\mathbf{z}$  is changing. This avoids the creation of interfaces at locations that are likely to change later in the inversion. Once the learned reference model has stabilized we can then incorporate the reference model into the smoothness term. It promotes the creation of interfaces, which also help to reach the smallness target misfit if it has not yet happened. This strategy and its effects are illustrated in the DC resistivity example (Section 5.2).

### 4.3 Convergence

At each iteration, the gradient-step for the geophysical objective function is similar to that in the Tikhonov formulation. This step is guaranteed to be a descent direction for the objective function. The objective function to be minimized, however, changes because the solution of the MAP-EM problem changes the reference model and smallness weights. Fortunately, the MAP-EM algorithm guarantees a decrease of the smallness term at each iteration. In other words the updated petrophysical distribution is guaranteed to be closer to the current geophysical model than the previous distribution.

Thanks to these properties, our algorithm usually reaches the geophysical target misfit at a rate comparable to the Tikhonov approach, and then requires only a few additional iterations to reach

the petrophysical target misfit. Each PGI example presented in the next Section 5 runs in a number of iterations comparable to the Tikhonov inversion while reaching both of our stopping criteria. This is a significant result as this means applying this framework does not necessarily end in a massive loss of efficiency.

## 5 EXAMPLES

In this section, we present two synthetic examples and a field example that illuminate important aspects of our framework. These examples are available as part of the SimPEG package on Github at <https://github.com/simpeg-research/Astic-2019-PGI> (Astic 2019). Readers are invited to use the notebooks and reproduce the results.

The first example is a non-linear inverse problem in electromagnetics. Magnetotelluric data (Ward & Hohmann 1988) are acquired over a layered-earth that has sharp and smooth features. We are provided with the true petrophysical distribution and the goal is to use it, along with the MT data, to find a solution that has the desired features. We present the results at the various steps of the algorithm and display the updates done to the initial petrophysical and geological models that are used to guide the geophysical inversion.

In the second example, a DC resistivity profile (Ward & Hohmann 1988) is acquired over two cylinders. We demonstrate how this framework performs when the GMM is not known a priori. We also use this example to show how prior geological information can be included into the inversion.

The third example is an airborne frequency-domain electromagnetics (FDEM) survey (Ward & Hohmann 1988) over a floodplain in Australia potentially contaminated by saline water. We highlight the influence of the choice of a reference model on the Tikhonov models. We then illustrate how the PGI approach can reduce ambiguity in the recovered features. This is achieved by constructing models in which the earth is assumed to have a certain number of distinct units. This sort of questions (what if there are  $c$  units?), crucial for interpretation, are otherwise difficult to ask in a traditional Tikhonov setting.

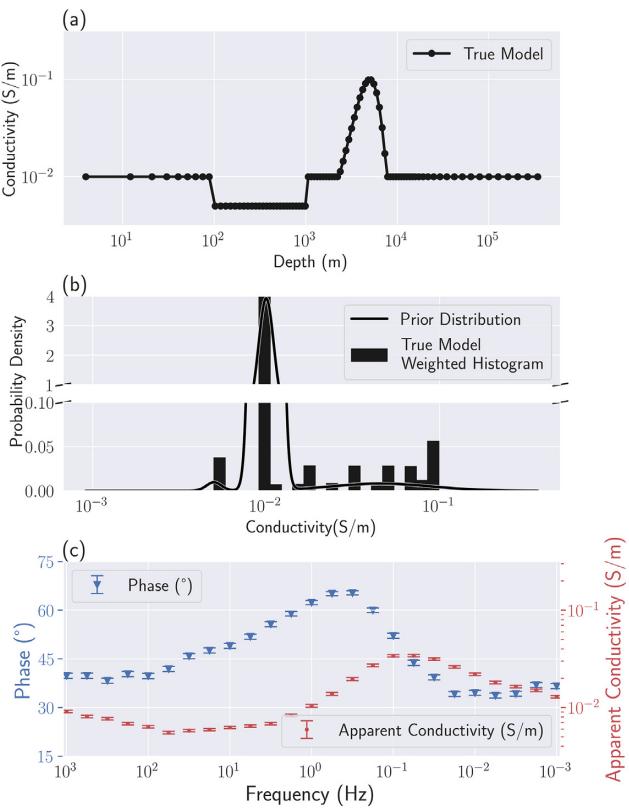
#### 5.1 MT 1-D, layered and smooth Earth model: demonstration of the framework

In this example, we demonstrate how to use the PGI framework to include petrophysical information. We first highlight the gains made in the recovered geophysical model. We then provide details about how the PGI iterations proceed.

##### 5.1.1 Setup

For this example, inspired by Kang *et al.* (2017), we want to recover a 1-D earth model that is made up a resistive layer and a smoothly varying unit embedded in a uniform background. The conductivity model, which is discretized onto a mesh with 89 cells, is shown in Fig. 7(a). The background unit has a conductivity of  $0.01 \text{ S m}^{-1}$  ( $100 \Omega\text{m}$ ). The resistive unit ( $5 \times 10^{-3} \text{ S m}^{-1}$  or  $200 \Omega\text{m}$ ) is 900 m thick and starts at a depth of 100 m. The smooth conductor, with a maximum conductivity of  $0.1 \text{ S m}^{-1}$  ( $10 \Omega\text{m}$ ), is located between 2300 and 7560 m depth. A log-scale, which is traditional for MT, is used.

In the MT experiment the electric and magnetic fields recorded at the earth's surface are combined to generate impedances. In this example, data are collected at 25 logarithmically spaced frequencies between  $10^{-3}$  and  $10^3$  Hz. This generates 50 data composed



**Figure 7.** Setup of the MT 1-D example: (a) true geophysical model  $\mathbf{m}_{\text{true}}$ . The dots coincide with the cell centres in the discretized model. (b) True model weighted histogram and prior petrophysical distribution  $\Theta_{\text{prior}}$ ; (c) MT data and uncertainties.

of the real and imaginary components of the impedance for each frequency. Unbiased Gaussian noise with a standard-deviation of 2 per cent of the data value is added. We represent those observed data in the form of apparent conductivity and phase in Fig. 7(c).

### 5.1.2 Initialization

We first describe our initialization for both the Tikhonov and PGI algorithms. For the initial geophysical model  $\mathbf{m}_0$ , we choose a uniform half-space with the same conductivity as the background unit. For the PGI approach the starting membership  $\mathbf{z}_0$  is set to the background unit everywhere. Both the Tikhonov and PGI approaches start with the same objective function; the models after the first iteration are thus identical. This helps for comparing results between the two approaches.

For the PGI approach we also need to provide a GMM to describe our prior petrophysical and geological information  $\Theta_{\text{prior}}$ . We use the exact information from the true earth model for the means and global proportions. For the variances, we use the true variance for the smooth unit. This unit is characterized by a wide distribution and thus it has a large variance. The background and the resistive unit are characterized by single values so their true variances are zero (Dirac distribution). We assign a standard deviation of  $0.1 \ln(\text{S m}^{-1})$  (or variance of  $0.01 [\ln(\text{S m}^{-1})]^2$ ) as an acceptable level of petrophysical noise for those two units. This prior petrophysical distribution can be visualized in Fig. 7(b). All the parameters were estimated by taking into account the ‘volume’ of the cells. This is necessary to make the GMM parameters independent of the discretization. The

histogram in Fig. 7(b) and all subsequent take into account the cell volumes and are thus referred to as weighted histograms.

We set all confidences in the prior petrophysical distribution  $\{\xi, \kappa, \nu\}$  to unity. Unity confidence parameters mean we have equal trust in the geophysical model and in the prior information when computing the new GMM parameters at each iteration of the petrophysical characterization. Each new estimate of the parameters, at each iteration of the MAP-EM algorithm (eqs 27, 32 and 34), is thus a simple unweighted average of the observed and priors. This allows us to see how well each parameter of the GMM is recovered without forcing or fixing its value in the inversion.

### 5.1.3 Comparison of the Tikhonov and PGI results

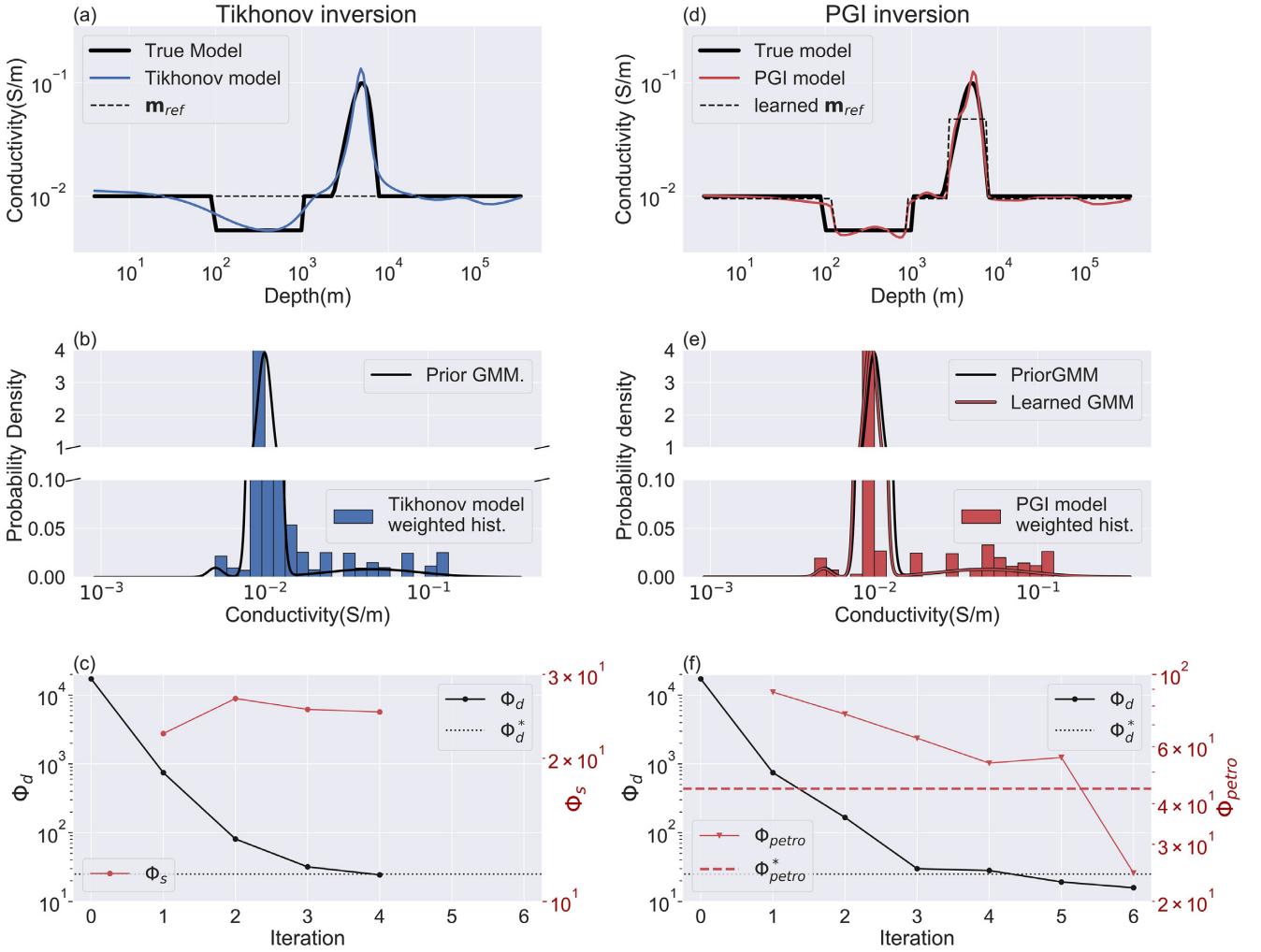
The results of the Tikhonov and PGI algorithms are shown in Fig. 8. The Tikhonov model is presented in Fig. 8(a) and the PGI model is shown in Fig. 8(d). The Tikhonov inversion is smooth everywhere, but character of units blockiness or smoothness are much better recovered in the PGI model. Note that both algorithms started from the same reference model. This reference model is kept constant in the Tikhonov approach whereas the PGI algorithm updates it, along with the smallness weights, at each iteration.

The PGI recovered petrophysical distribution (Fig. 8e) is also much closer to the prior and true distributions, compared to the Tikhonov result (Fig. 8b). The means, variances and proportions of each unit are well recovered. Despite the fact that we used only confidences of unity in the prior GMM parameters, we were able to recover a learned petrophysical distribution that was close to the true distribution. The recovered variance of the background is  $(7.2 \times 10^{-3} [\ln(\text{S m}^{-1})]^2)$ . This is less than our prior value of  $0.01 [\ln(\text{S m}^{-1})]^2$ . This means that we recovered a background conductivity with a smaller variance than the one prescribed by the prior.

In addition to being closer to the ground truth, our inversion procedure converges in a similar manner as the Tikhonov approach (Figs 8c and f). The PGI method took only two additional iterations, compared to the Tikhonov approach, to reach the geophysical and petrophysical target misfits (shown respectively in black and red in Fig. 8f). They are reached respectively at the fifth and sixth iterations. Those two Figs 8(c) and (f) also highlight that the smallness is now a term to be minimized in the PGI. In Tikhonov we are unconcerned about the value of  $\Phi_s$  and it usually increases when  $\beta$  decreases. In PGI,  $\Phi_s$  measures how well the petrophysical and geological information are recovered.

### 5.1.4 Step-by-step of the PGI iterations

The iterations of the PGI are shown in Fig. 9. Figs 9(a) and (b) show our initialization. The initial and reference model is a half-space with a conductivity equal to the background conductivity. The prior and initial petrophysical distribution  $\Theta_{\text{prior}}$  was described in the initialization section for this example. The first iteration (Figs 9c and d) is the same as the first Tikhonov iteration and the current recovered geophysical model is fairly smooth. In the Tikhonov inversion, the reference model remains the same throughout the inversion but in the PGI, the reference model and smallness weights are updated at the end of each iteration. The updated reference model is determined from the learned petrophysical distribution applied to the current geophysical model, as described in eqs (21)–(23). The recovered petrophysical distribution at this first iteration already distinguishes



**Figure 8.** MT 1-D example results and comparison: panels (a), (b) and (c) show the recovered geophysical model, its histogram and the convergence curves, respectively, for the Tikhonov approach. Panels (d), (e) (f) shows the same plots for the PGI approach.

three units but displays higher variances than the prior. The conductivity contrasts are underestimated but the overall geological classification is correct. The recovered geophysical model after the second step (Figs 9e and f), with the updated reference model and smallness weights from the previous step, starts to display the desired features both spatially and in its physical property distribution; we see a sharper resistive unit and smoother conductive unit. The geophysical target misfit is reached at iteration 5 (Figs 9k and l) and the parameter  $\alpha_s$  is increased. In one further step (Figs 9m and n) both the geophysical and petrophysical target misfits are reached. The last iteration successfully clustered the model values while conserving  $\Phi_d \leq \Phi_d^*$ . The sharp contrast of the resistive unit is well recovered as well as the smoothness of the conductive unit. The learned petrophysical distribution is very close to the prior distribution. It also has a lower variance for the background unit compared to its prior value, and thus is closer to the truth.

## 5.2 DC 2-D, buried cylinders: Working with missing petrophysical information and adding geological information

In this example we illustrate the performance of our framework when no physical property mean values are available, and compare

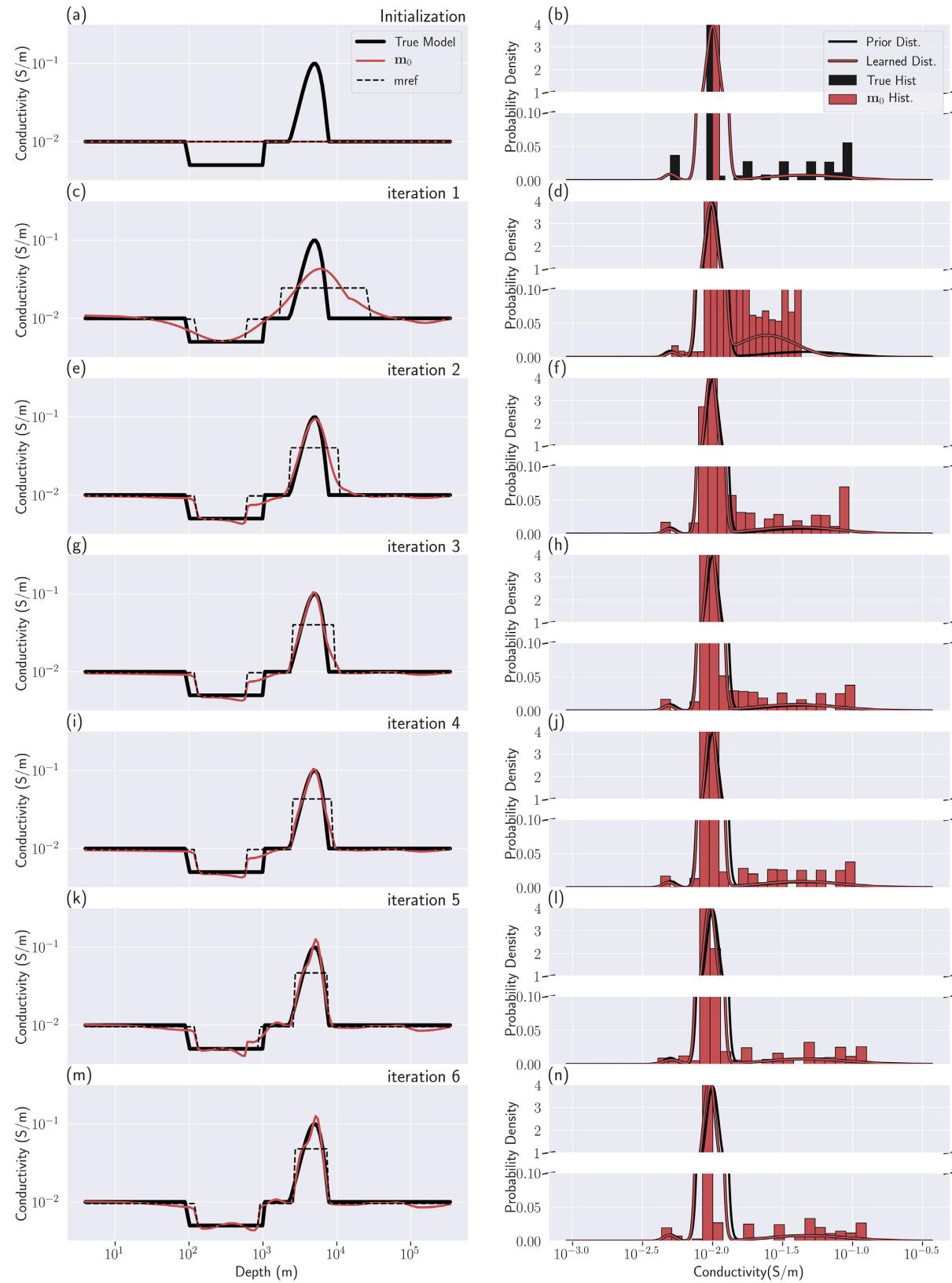
it to the result when full petrophysical information is available. We then show how geological information, such as depth and thickness, can be incorporated.

### 5.2.1 Setup

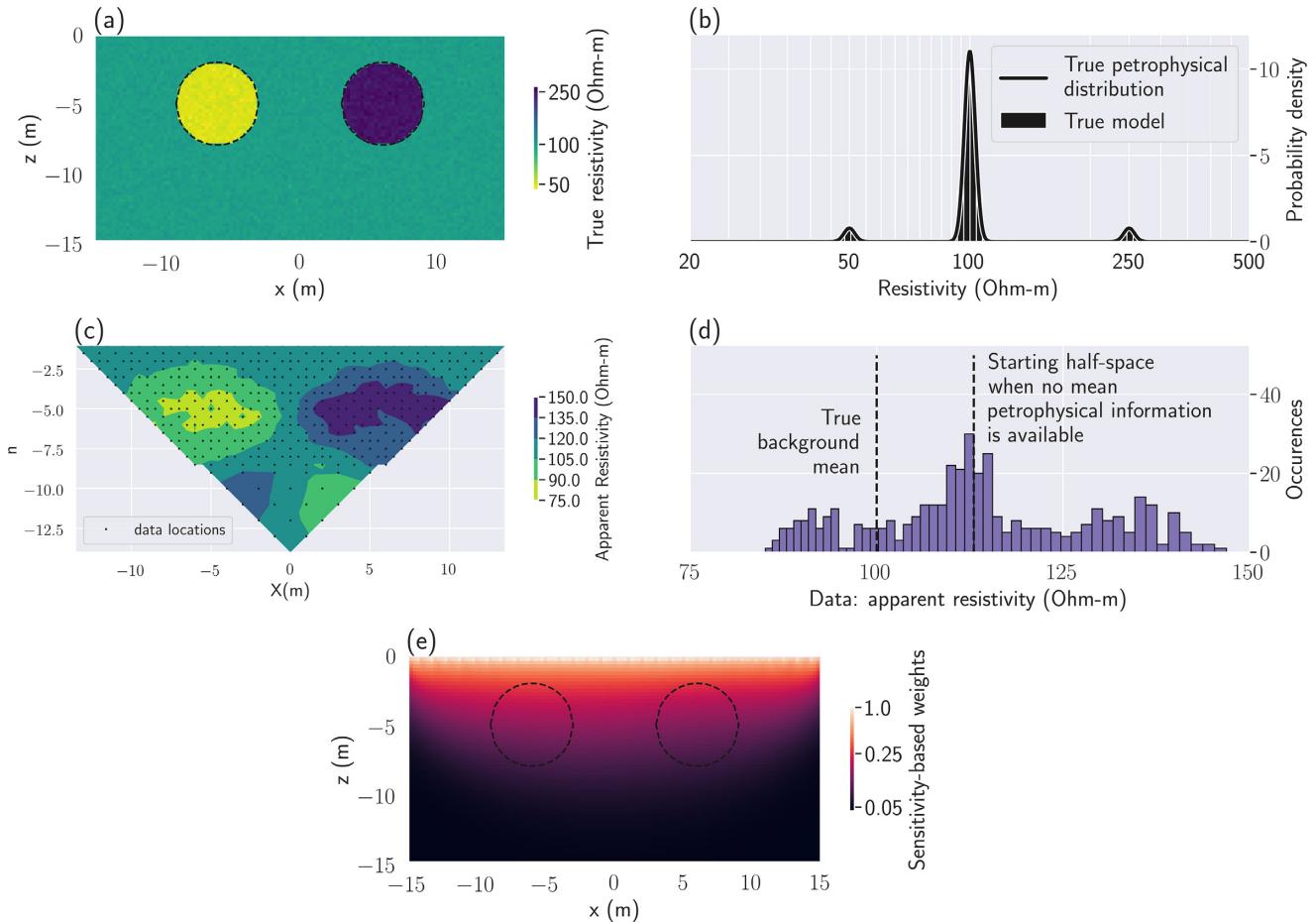
We apply our procedure to a 2.5-D DC resistivity problem (2-D geology but 3-D sources) to recover two cylindrical units, one conductive (with a mean resistivity of 50  $\Omega\text{m}$ ), one resistive (with a mean resistivity of 250  $\Omega\text{m}$ ) embedded in a background unit with a mean resistivity of 100  $\Omega\text{m}$  (see Fig. 10a). To make the model slightly more geologically realistic, we added Gaussian noise to the log-conductivity model. The noise had zero mean and a variance of  $1 \cdot 10^{-3} [\ln (\text{S m}^{-1})]^2$  (see Fig. 10b).

Our survey is a dipole-dipole with electrode separation of 1 m and 2 m (Fig. 10c). The maximum dipole separation is set to  $n = 16$ . We simulate a total of 419 resistance measurements. Unbiased Gaussian noise, with a standard-deviation of 2% of the original value, is added to the forwarded geophysical data. A histogram of the apparent resistivity data is shown in Fig. 10(d).

For the inversion we limit our active cells to the region covered by the survey data as shown in Fig. 10(a). The active mesh is composed of 7021 cells. For the geophysical data misfit uncertainties, we used



**Figure 9.** Iterations of the PGI MT 1-D example.



**Figure 10.** Setup of the DC resistivity example. (a) True conductivity. (b) True petrophysical distribution. (c) DC data to be inverted. (d) Histogram of the measured apparent resistivity. The best-fitting half-space ( $112 \Omega\text{m}$ ) is indicated by the vertical dashed line. (e) Sensitivity weights  $w$  used for all inversions.

a floor value of  $10^{-4} V/I$  in addition to the 2% noise standard deviation.

For all of the inversions we used a sensitivity weighting that is similar to that developed in Mehane et al. (2005). This is used in place of the usual smoothing weights at the surface to compensate for the high sensitivities near electrodes. Those weights are shown in Fig. 10(e). We use  $\Phi_d$  and  $\Phi_{\text{petro}}$  as stopping criteria. To illustrate functionality, we also incorporated the learned reference model in the smoothness term of the geophysical objective function. This was done towards the end of the inversion, after the reference model was stabilized and  $\Phi_d^*$  was reached (see Algorithm 1, step 6).

To set benchmarks for the inversion results, we first carry out a Tikhonov inversion and a PGI with full petrophysical information. We then run a PGI without providing any information about the physical property mean values or the proportions. We finally run another PGI, still without means information, but with added geological information included through the use of local proportion weights in the GMM. All inversions start with the same geophysical objective function.

### 5.2.2 Tikhonov inversion

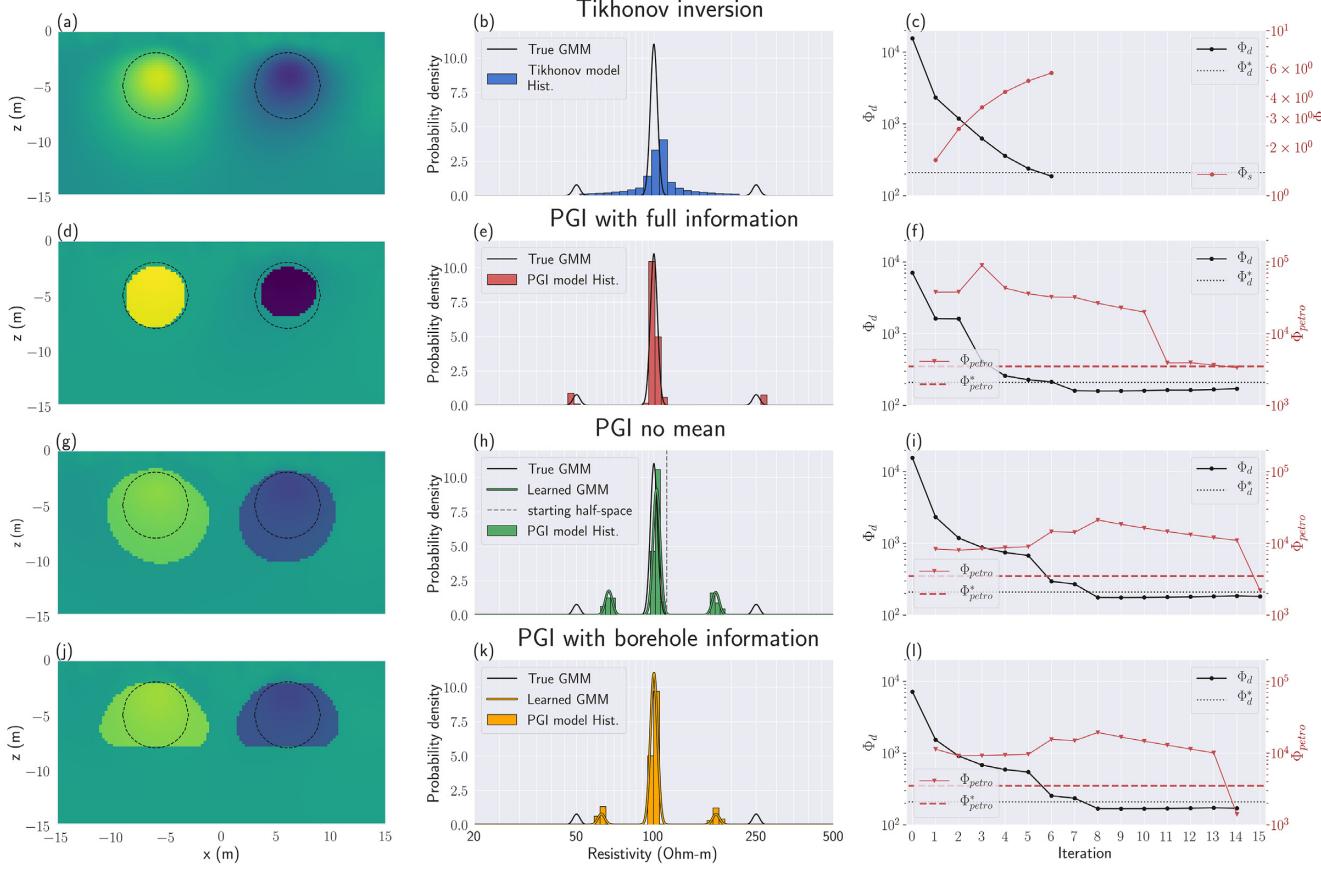
The Tikhonov inversion starts from the best fitting half-space of  $112 \Omega\text{m}$ , instead of the true background value of  $100 \Omega\text{m}$  (see Fig. 10d). The inversion produces the result presented in Fig. 11(a).

The two bodies are detected but their edges are smoothed. This smoothness is also visible in the histogram of the model that is characterized by continuous values centred on the starting half-space value (Fig. 11b). The conductivity values do not attain the true electrical conductivity of the anomalies. The geophysical target misfit is reached in six iterations (Fig. 11c). The parameter  $\beta$  was cooled at each iteration and the smallness term kept increasing; this is expected in the Tikhonov approach.

### 5.2.3 PGI with full petrophysical information

In this example we first want to establish a benchmark result by using the true petrophysical distribution with global proportions. We fix the petrophysical distribution at the true one (Fig. 10b) by setting all our confidences in the petrophysical prior to infinity. Thus  $\Theta = \Theta_{\text{prior}} = \Theta_{\text{true}}$  is fixed for all iterations. We start from a half-space with the true mean background value.

This benchmark inversion recovers the two cylinders (Fig. 11d) quite well. The locations of the cylinders, their outer boundaries and their conductivities are in reasonable agreement with the true model. The geophysical model also satisfactorily fits the true petrophysical distribution being imposed (Fig. 11e). The geophysical target misfit is reached after seven iterations (Fig. 11f). An additional seven iterations is needed to reach the petrophysical target misfit, while keeping  $\Phi_d \leq \Phi_d^*$ . The two increases of  $\Phi_{\text{petro}}$  seen at iteration



**Figure 11.** DC inversion results. The color scale is the same as for the true model in Fig. 10(a). Panels (a), (b) and (c) present the recovered model, histogram and convergence curves, respectively, for the Tikhonov approach. Panels (d), (e) and (f) show the same plots for PGI, when the true petrophysical distribution was provided. Panels (g), (h) and (i) show the results for PGI, when no information on the proportions or the physical property means of the units were provided. Panels (j), (k) and (l) show how PGI, with no mean information, can be improved by adding information about the depth and thickness of the cylinders.

3 and 7 correspond to  $\beta$ -cooling. This was necessary since the geophysical misfit was not decreasing enough (see algorithm 1, step 6). The important decrease seen at iteration 11 corresponds to the inclusion of the learned reference model in the smoothness. This highlights the importance of incorporating a discontinuous reference model into the smoothness terms so that large gradients in the model are not penalized. This allowed for the final model to have sharp boundaries which is consistent with the true model.

#### 5.2.4 PGI with no physical property means information

We now consider the situation where no prior information is known, except the expected number of units (three in this example). We can still use our framework but we turn off the prior on the means and proportions. The confidences  $\{\zeta, \kappa\}$  are thus set to zero (no prior information on the means or proportions). We used the true variances with infinite confidences  $\nu$  in them. We thus have specified the petrophysical noise levels, which regulate how much each unit has to cluster around its unknown mean. This is analogous to choosing an appropriate geophysical data noise level. Our inversion starts from the best fitting half-space, as in the Tikhonov inversion, since the true background value is now considered to be unknown.

The recovered geophysical model still displays a clustered aspect with structures close to the ground truth (Fig. 11g). The volumes of the two cylindrical units are overestimated. The histogram and

model show three distinct units. Without the true means information, the recovered means of the cylinders are slightly shifted towards the background value (Fig. 11h). Interestingly the recovered mean of the background is closer to the true mean than the initial half-space model. The number of iterations to reach both target misfits is comparable to the PGI example that used full petrophysical information.

#### 5.2.5 PGI with geological information

Our next goal in this example is to illustrate how geological information might be included. It is motivated by the fact that, in the two previous PGI results (using full and no petrophysical information), the cylindrical units display various size anomalies compared to the ground-truth and extend to different depths. Suppose information about the anomalous zone's depth and thickness are provided. This additional information can be added to the inversion through local proportion parameters  $\mathcal{P}(\mathbf{z})$ , by making them vary with depth. At depths where the anomalous bodies are expected, we set equal prior expectation of encountering any of the three units (so here  $\pi = 1/3$  for all three units for locations at depth between 2 and 8 m). At depths where only the background is expected, we set  $\mathcal{P}(\mathbf{z} = \text{background}) = 1$  and  $\mathcal{P}(\mathbf{z} = \text{cylinders}) = 0$ . Those proportions are summarized in Table 1. We used this prior information in a new PGI, using the petrophysical distribution recovered from the PGI with no mean information as  $\Theta_{\text{prior}}$ .

**Table 1.** Proportions of the GMM  $\mathcal{P}(\mathbf{z})$  as a function of depth to include geological information. Bckgrd: background; cond.: conductive unit; res.: resistive unit.

Depth range (m)	$\mathcal{P}(\mathbf{z} = \text{bckgrd})$	$\mathcal{P}(\mathbf{z} = \text{cond.})$	$\mathcal{P}(\mathbf{z} = \text{res.})$
$\leq 2$	1	0	0
2–8	1/3	1/3	1/3
$\geq 8$	1	0	0

The recovered model (Fig. 11j) is similar to the PGI with no petrophysical information except the anomalies are now restricted to the right depths. Comparing the two learned distributions, with and without depth information, we notice that the recovered conductivity values from the PGI model with geological information are closer to the ground-truth than the conductivities recovered from the PGI with no petrophysical or geological information (Fig. 11 k). The convergence of the algorithm is similar to the previous PGIs (Fig. 11 l).

### 5.3 Using PGI to reduce ambiguity: a field example

In any inversion the final result is affected by many terms and parameters in the objective function, but of particular interest here is the choice of reference model. In a Tikhonov inversion starting from different half-space reference models can lead to different solutions, and this complicates the interpretation. Some of the ambiguity might be reduced if more information is incorporated. The PGI approach allows us to add information, such as a desired number of geological units, that are otherwise difficult to use in a Tikhonov setting. Adding a constraint on the number of units we wish to recover, without incorporating other extensive geological or petrophysical information, can be enough to reduce the realm of possible models.

We illustrate this on a field frequency-domain EM data set (Ward & Hohmann 1988) for saline water delineation. We begin by emphasizing the discrepancies seen in Tikhonov models using different half-space reference models. We then show that by using PGI and assuming an expected number of distinct units, that starting from different reference models leads to similar inversion results that have consistent petrophysical distributions and simplified structures. The lack of dependence upon the initial and initial reference models helps to build confidence in the final images.

#### 5.3.1 Setup

The Bookpurnong irrigation District is part of the Riverland region of South Australia along the Murray river. The irrigation on the highland river bank has lead to the salinization of the floodplain soil and threatens to make the area inhospitable for vegetation (Fig. 12). The key issue is to determine if (and where) the fresh water river is charging the aquifer (called ‘losing stream’, this is the healthy scenario), or if the saline aquifer is charging the river, and thus contaminating the floodplain soil (‘gaining stream’ scenario). To help the work of hydrogeologists, various frequency and time-domain electromagnetics surveys have been conducted to characterize saline zones which are diagnosed by an increase in electrical conductivity. High conductivities close to the surface will indicate a ‘gaining stream’ scenario while low conductivities will be a sign of a ‘losing stream’.

The Bookpurnong case study has been studied in previous publications (Viezzoli *et al.* 2009, 2010; Yang 2017). Here we focus on the RESOLVE frequency domain data set, flown in 2008, which covers an area of approximately 6650 m by 2400 m.

At each sounding, the RESOLVE system measures the real and imaginary parts of the induced magnetic field for five frequencies ranging from 382 Hz to 130.1 kHz. The survey consists of 1022 soundings, resulting in a total number of data of 10 220. For the geophysical noise levels, we use a standard-deviation of 10 per cent of the absolute value of the datum and add a 20 ppm floor value; this is consistent with previous studies.

#### 5.3.2 Inversion setup

To invert the RESOLVE data set, we choose a laterally constrained 1-D approach similar to Viezzoli *et al.* (2008) using the EM1D module of the SimPEG package (Heagy *et al.* 2017; Kang *et al.* 2018). Each sounding is inverted in 1-D using a mesh with 19 layers that extends to a depth of 187 m. The cells increase in size from 2.5 m at the surface to 25 m for the last layer. We use the thickness of the layers as weights in the regularization, which relates to an integral formulation of the regularizer (Oldenburg & Li 2005).

For each Tikhonov and PGI approach, we run two separate inversions using  $0.01 \text{ S m}^{-1}$  ( $100 \Omega\text{m}$ ) and  $1 \text{ S m}^{-1}$  ( $1 \Omega\text{m}$ ) models, respectively, as reference models (thus four inversions are carried out). The reference models are also used as initial models  $\mathbf{m}_0$ . The reference model stays the same throughout the Tikhonov inversions while the PGI algorithm updates it.

#### 5.3.3 Tikhonov inversions

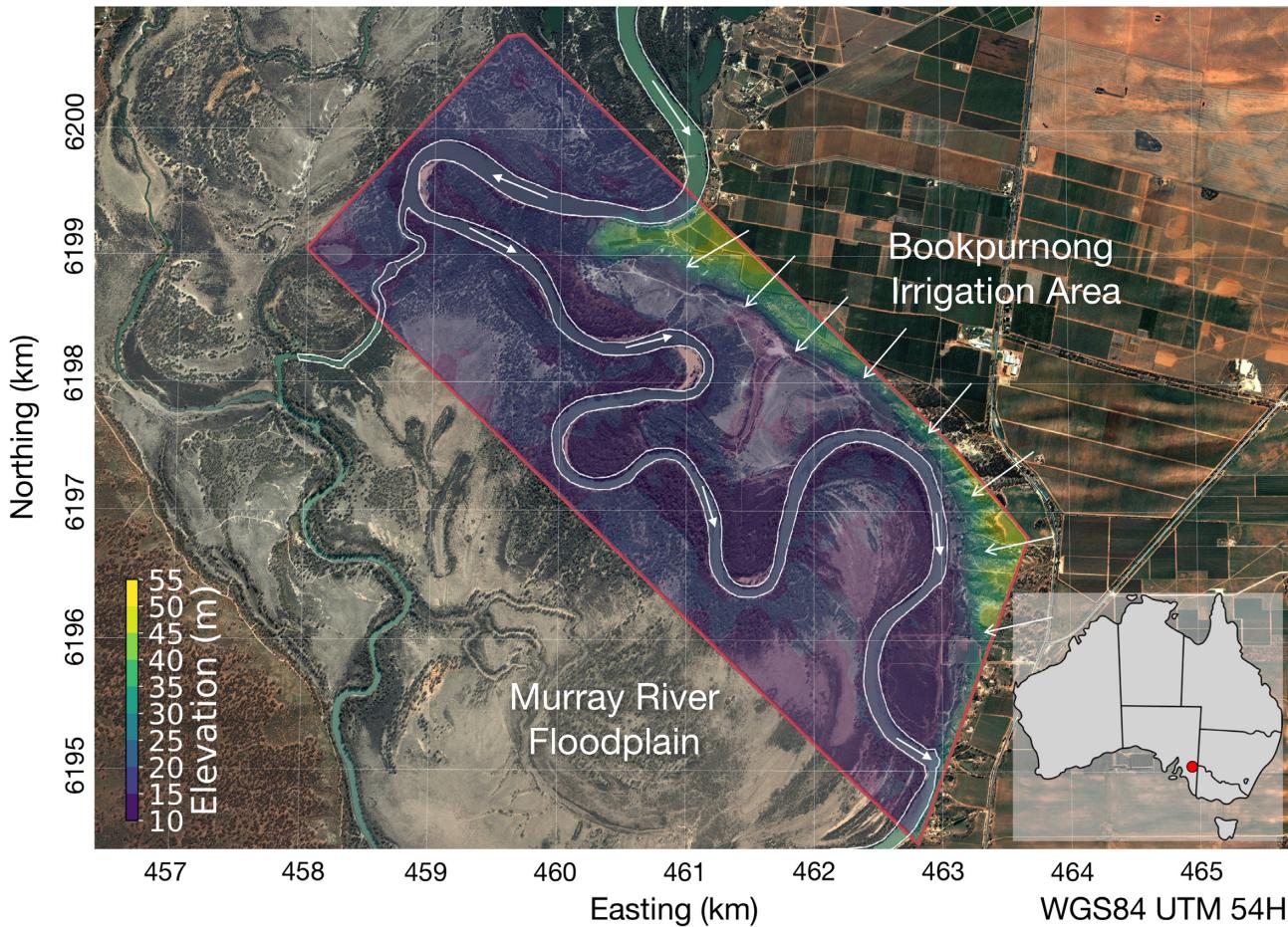
The results of the Tikhonov inversions using two different reference models are summarized in Fig. 13. Both Tikhonov models (Figs 13a and d) detect the river at the surface as a resistive feature; it is also visible where the cross-section crosses the river near 459, 459.8, 460.5, 461, 461.8 and 462.5 km on Figs 13(b) and (e). Both models also pick up highly conductive features at depth and the laterally constrained regularization helps ensure horizontal continuity. The ranges of conductivity are similar to those observed in boreholes (Holland *et al.* 2008).

In the northwest part of the area, the recovered conductivity models are consistent and transition smoothly from resistive at the surface to conductive at depth. In the southeast part of the area, more discrepancies in the two models are visible (see Section 5.3.5 for a quantification of the discrepancies). On the plan views (Figs 13a and d), the overall recovered conductivities in the biggest bend in the river are an order of magnitude different between the two inversions. Inconsistent pocket-like structures are seen in both inversions in the cross-sections (Figs 13b and e). Both cross-sections however show that the conductive layer is closer to the surface in the southeast than in the northwest.

#### 5.3.4 PGI

We now apply the PGI approach without providing specific petrophysical information; only the number of units is specified. We show that similar geophysical models and petrophysical distributions are obtained even though the inversions are carried out with different initial reference models.

As we did for the Tikhonov inversions, we now run two inversions starting from uniform  $0.01 \text{ S m}^{-1}$  ( $100 \Omega\text{m}$ ) and  $1 \text{ S m}^{-1}$  ( $1 \Omega\text{m}$ ) models, respectively, used as initial and initial reference models. For the conductive and resistive reference model cases, respectively, both the Tikhonov and PGI approaches start with the same objective function; the models after the first iteration are thus identical.



**Figure 12.** Bookpurnong survey area. The elevation of the survey area is represented as an overlaid colormap. The white arrows represent the water flow, both of the river and from the irrigation zone. Geographic system: WGS 84 / UTM 54S. Aerial images from Google Earth, ©2019 DigitalGlobe.

For the prior petrophysical distribution, we do not specify any prior value for the mean conductivity values nor the proportions; their respective confidence parameters are thus set to zero. We choose variances of  $0.1[\ln(\text{S m}^{-1})]^2$ . We set the confidences in the prior covariances to unity. Our goal with this choice is to recover a model with distinct units but still reasonably smooth.

A critical point here is to choose the number of distinct units we want to recover. We assume that there are three units, as this matches our interpretation goal of distinguishing freshwater from salinized water, with an additional transition zone. This adds an additional constraint on the models we want to recover.

The PGI results are compiled in Fig. 14. All PGI have reached their geophysical and petrophysical misfits. The recovered models (Figs 14a and d) are more consistent with each other than those in the Tikhonov case; we will discuss this later in Section 5.3.5.

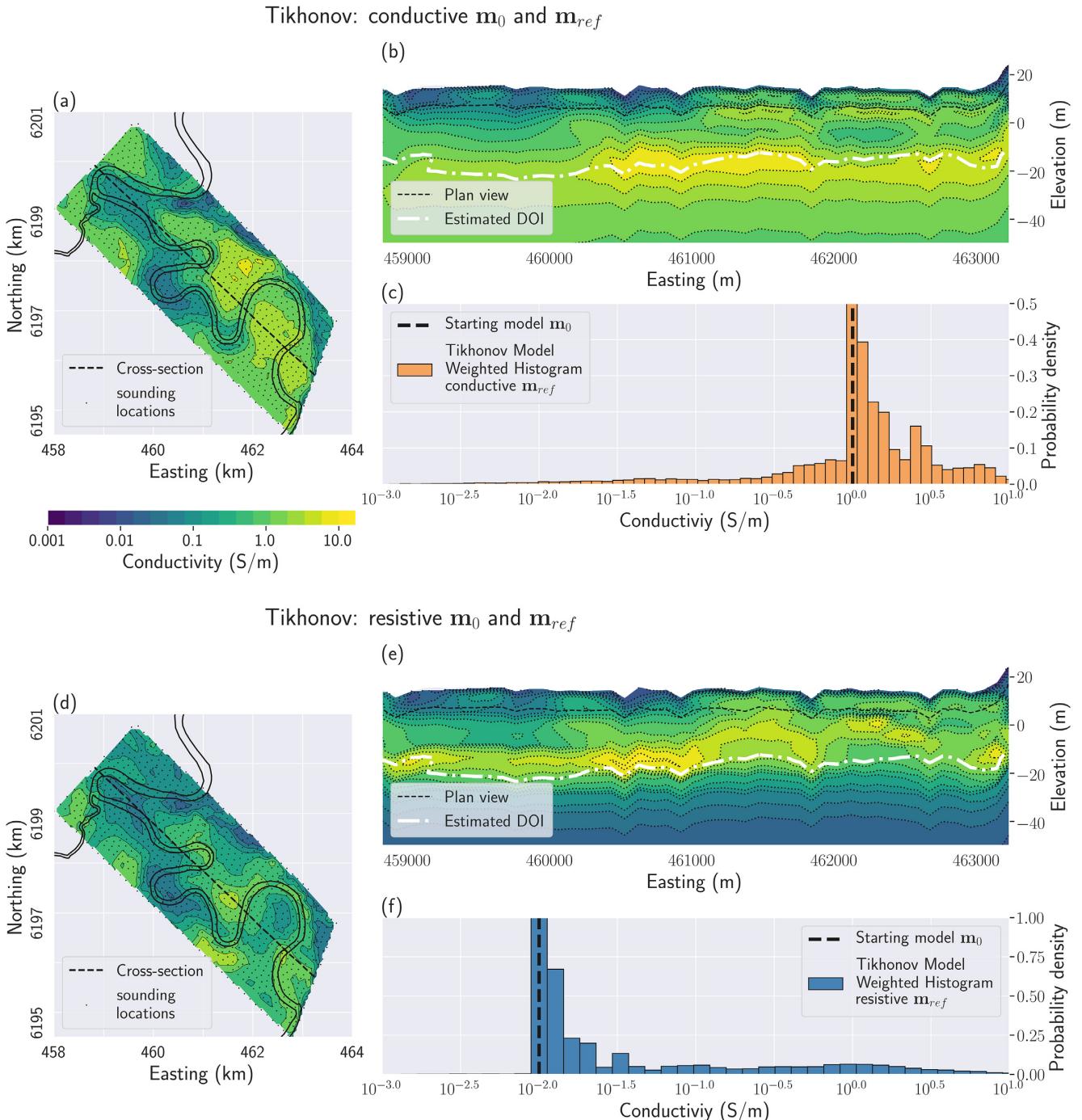
Counterintuitively, by clustering the histograms we actually smoothed the resulting model by forcing more homogeneity inside each distinct unit. Neither cross-section (Figs 14b and e) shows the pocket-like structure in the southeast end of the profile that were observed in the Tikhonov inversions; it indicates that those features were not robust. The high conductivity layer is closer to the surface in the southeast than in the northwest in both models.

The histograms and recovered petrophysical distributions (Figs 14c and f) display several interesting characteristics. First, we recover three distinct units that can be interpreted as fresh, transition and saline zones. The learned means of the clusters are very

similar between the two inversions. We notice that in the case of the resistive initial models, the resistive unit conductivity has been corrected to about  $0.017 \text{ S m}^{-1}$  ( $60 \Omega\text{m}$ ). The same phenomenon happens for the conductive initial models case, with a corrected conductive unit conductivity value of  $1.35 \text{ S m}^{-1}$  ( $0.74 \Omega\text{m}$ ). This behavior was noticed in the DC example (Section 5.2) when starting from a biased background value. The major difference between the GMM for the two models is the variance and proportion of the most resistive and most conductive units. This is due to the volume of the region in the padding zones that are outside the region of influence of the data. The values of conductivity in those regions stays close the cluster mean which is closest to the conductivity of the initial reference model.

### 5.3.5 PGI - Tikhonov comparison

In addition to the removal of unnecessary structure mentioned earlier, the models recovered with the PGI approach display more consistent subsurfaces features than do the Tikhonov models. To quantify the similarities between models within the same methodology but with different initial and reference models, we adopt an approach analogous to the depth of investigation (DOI) estimation presented in Oldenburg & Li (1999). For the two Tikhonov and the two PGI conductivity models, respectively, which for clarity in the



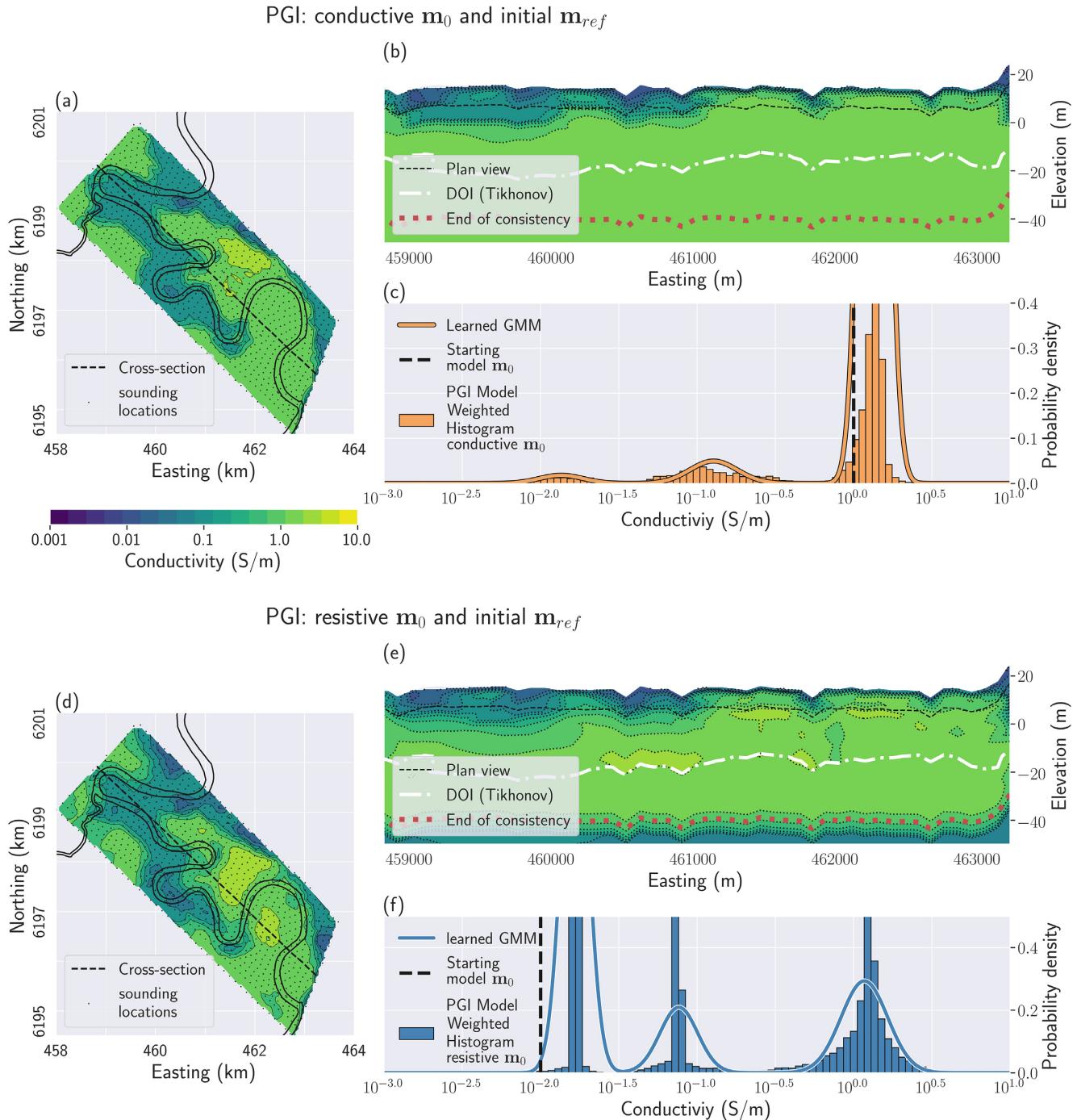
**Figure 13.** Tikhonov results with two different reference models. All figures share the same color scale for conductivity. Panels (a), (b) and (c) show the results starting from a conductive half-space. Panels (d), (e) and (f) show the result starting from a resistive half-space. Panels (a) and (d) display the recovered geophysical models on a plan view at a depth of 8.5 m. Figs (b) and (e) display a cross-section of the recovered geophysical models. Depth of investigation (DOI) evaluated using the method presented in Oldenburg & Li (1999). Panels (c) and (f) display the histogram of the recovered model for each inversion.

notation are referred to in the following equation as electrical resistivity  $\rho_1$  and  $\rho_2$ , we computed their maximum ratio  $r$  everywhere:

$$r = \max \left( \frac{\rho_1}{\rho_2}, \frac{\rho_2}{\rho_1} \right) (1 \leq r \leq \infty). \quad (42)$$

A ratio  $r$  of unity indicates similar values of conductivity. The higher the value of  $r$ , the more the two models differ. Fig. 15 shows the result of this two-by-two comparison. The two PGI models are

quite consistent, with ratio  $r$  values close to unity almost everywhere, except in a small area in the northwest. On the contrary the Tikhonov inversions show significant discrepancies, even in the near-surface, with  $r$  values greater than 2 in most of the survey area. This is especially visible near the important bend in the river in the southeast where discrepancies of more than an order of magnitude ( $r \geq 10$ ) are observed (dark areas in Fig. 15). Fig. 15(j) highlights when the Tikhonov inversions reached their DOI and where the final



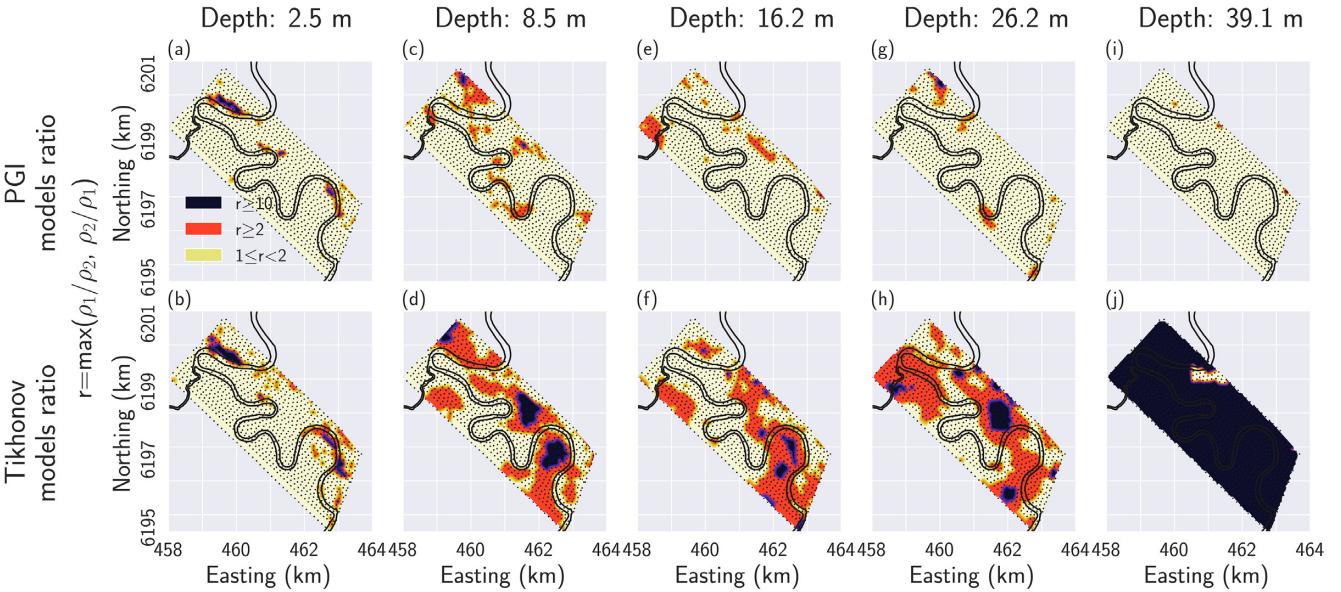
**Figure 14.** PGI results with two different initial and initial reference models. All figures share the same color scale for conductivity. Panels (a), (b) and (c) show the results starting from a conductive half-space. Panels (d), (e) and (f) show the results starting from a resistive half-space. Panels (a) and (d) display the recovered geophysical models on a plan view at a depth of 8.5 m. Panels (b) and (e) display a cross-section of the recovered geophysical models. Panels (c) and (f) display the histogram of the recovered model and the learned petrophysical distribution for each inversion.

model is controlled mostly by the regularization. It is worth noticing that the PGI models are still consistent past the DOI estimated from the Tikhonov inversions (Figs 14b and e).

#### 5.3.6 Field example summary

We demonstrated that even in the near-surface above the depth of investigation the Tikhonov inversions had different structures

when different initial and reference models were used. With the PGI approach, the two inversion lead to similar near-surface models that have consistent petrophysical distributions. The removal of structures, evidently not required by the data, and the simplification of the geology, provides a result that is more appealing to interpret geologically. Making a reasonable guess about the variances and the number of units were the crucial additional information required to achieve this. All four models point to a ‘losing stream’ scenario



**Figure 15.** Consistency comparison between the two Tikhonov and the two PGI models, starting with conductive and resistive initial and reference models. The first row displays the ratio  $r$  of the two PGI conductivity models at different depths. The bottom displays the same ratio for the two Tikhonov conductivity models. Depths of the plan views increase from left to right. Pale yellow means the two conductivity models differ by less than a factor 2. Orange highlights areas where the recovered conductivities differ by more than a factor 2. Dark purple emphasizes area where the two recovered conductivity models differ by more than a factor 10.

in the northwest while three (both PGI models and the Tikhonov inversion with a conductive reference model) indicate a ‘gaining stream’ in the southeast. This is coherent with the conclusions given in Viezzoli *et al.* (2009, 2010) and Yang (2017). To further extend our analysis, ground-truthing the depth of the saturation salinization zone and adding it into the PGI as geological information, or adding petrophysical measurements, would be required.

## 6 DISCUSSION

We have presented a framework for carrying out a petrophysically and geologically guided inversion which allows the user to recover a quasi-geological rock model (Li *et al.* 2019) and a petrophysical distribution in addition to the geophysical model. The framework has three main modules, each of which is formulated as a MAP estimate. The framework is flexible and goals, achieved by others and using different approaches, can be incorporated while the reverse is not true.

Carrying out an inversion that includes three separate subproblems that communicate with each other, and for which two target misfits are to be achieved, is challenging. To address this we have provided a pseudocode as well as a flowchart in Appendix C. The important part of the procedure is the initialization step where the user must specify what is known, how to incorporate it, and what they want as an end product. The examples explore those various aspects as well as the convergence of the PGI approach.

The crucial component of our framework is to provide a different role for the smallness component of the regularization term in the geophysical objective function. An evolving reference model, tied to different units in a GMM, ultimately allows this term to be viewed as a quantitative metric for determining how well the petrophysical and geological data are fit. This has many advantages. The inverse problem is now solved by finding a model  $\mathbf{m}$  and a reference model  $\mathbf{m}_{\text{ref}}$ , along with the smallness weights, such that the

geophysical, petrophysical and geological data are fit. This complicates the inversion but the procedure is greatly simplified compared to adding additional terms, along with their adjustable hyperparameters, in the regularization. This also allows this research to be incorporated with other developments done using the Tikhonov approach. For example when sharp contrasts are expected, such as in the DC example developed in Section 5.2, the effects of using sparse norms for the model gradients in the smoothness term (Fournier & Oldenburg 2019) are worth exploring. This would allow the combined use of petrophysical and structural data to guide the inversion.

Our solution is achieved by first focusing on the geophysical data misfit and solving the optimization problem at successively smaller values of the global trade-off parameter  $\beta$ . In some cases, finding an acceptable model to the geophysical data also makes the petrophysical misfit sufficiently small. If it doesn’t, then we adjust  $\alpha_s$  which controls the relative weight between the smallness and the other components of the objective function. So far our heuristic approach has worked satisfactorily but searching simultaneously for optimal  $\beta$  and  $\alpha_s$  values such that both  $\Phi_d(\mathbf{m}) \simeq \Phi_d^*$  and  $\Phi_{\text{petro}}(\mathbf{m}) \simeq \Phi_{\text{petro}}^*$  is a more advanced optimization process and deserves further research. For example, Sun & Li (2018) is entirely dedicated to the problem of refining the values of  $\beta$  and  $\alpha_s$  within the FCM-guided inversion approach (Sun & Li 2015), and only for magnetic linear problems.

Further work is also required about the best way to evaluate the petrophysical target misfit. In our framework we minimize  $\Phi_s$  which is a weighted petrophysical misfit. Those weights are usually related to the physics of the survey and they are often necessary for solving inverse problems where sensitivities are greatly varying (e.g. in potential field data). Since we know the weights, we can evaluate  $E[\Phi_s]$  and use that as a fitting criterion. Another option is to evaluate  $\Phi_{\text{petro}}$ ; this amounts to evaluating  $\Phi_s$  without weights. Both have worked for us, but working with  $\Phi_{\text{petro}}$  has constantly produced good results and it is appealing to have a stopping criterion

ties only to the petrophysical and geological characteristics of the model.

Another key component of our framework is the possibility for the algorithm to learn all the parameters of a suitable GMM distribution when little to no information is known *a priori*. The MAP-EM algorithm has been an essential element in achieving this goal. The minimum requirement is to provide an expected number of distinct units and a prior petrophysical noise level for each unit.

The MAP-EM algorithm with semi-conjugate priors formalizes the update of the means done in Sun & Li (2015) and generalizes it to the other GMM parameters. An important aspect of our formulation of the MAP-EM algorithm is the incorporation of the volume elements into the evaluation of the GMM parameters, rather than just working with cell-counts. Working with volumetric values ensures that our recovered GMM parameters are independent of the discretization. The estimation of the GMM parameters includes a prior weighted by our confidences in this prior, in a similar sense that we have a confidence  $\beta$  in the geophysical model prior during the geophysical inversion. So far values of unity or above for the confidences in the variances have worked well whether or not we knew the true petrophysical distribution. Confidences in the means and proportions appear dependent of the quality of the prior knowledge. Effects of their values is to be further investigated, as well as if cooling or warming those parameters, similar to what is done with  $\beta$ , can have an impact on the recovery or convergence.

## 7 CONCLUSIONS

We have developed a framework for carrying out petrophysically and geologically guided inversions using a dynamic Gaussian mixture model prior. Importantly we achieve our goal of incorporating both geological and petrophysical information in geophysical inversions without including an additional term in the objective function. Rather, we update the reference model and the smallness weighting matrix at each iteration of the geophysical model through the optimization of posterior probability distributions. This allows our work to be compatible with previous Tikhonov approaches and readily adapted to existing inversion codes. Our flexible prior formulation allows us to refine the petrophysical model as part of the iterative process. Our inversion continues until the geophysical, geological and petrophysical data are fitted. For that purpose, we have defined a target misfit to measure how well the recovered geophysical model fits the petrophysical and geological distribution. We have presented a suite of synthetic and field examples to illustrate important aspects of our framework, especially in demonstrating how prior petrophysical and geological information is incorporated. We have also shown that detailed knowledge of the petrophysical or geological properties is not required to make significant gains in the recovered geophysical models. Our examples deal with a single physical property but our framework has been designed to carry out joint inversion of different types of geophysical data, as illustrated in Astic & Oldenburg (2018); this topic will be addressed in a follow-up paper.

Finally, we comment upon how important the development of open-source software packages has been for our research. It has allowed us to benefit from the work carried out by others. The development of the MAP-EM algorithm has been facilitated by the MLE code available through the open source Scikit-Learn project, a library of tools for machine-learning (Pedregosa *et al.* 2011). The successful application of our methodology to various geophysical survey methods, such as potential fields, DC, MT or FDEM, has

been made possible because of the interconnectivity of SimPEG, an open source package to carry out geophysical inversions (Cockett *et al.* 2015; Heagy *et al.* 2017; Kang *et al.* 2018). To ensure reproducibility of the results in this paper, and to contribute to the development of the open source community, we are making the examples presented in this paper available online (Astin 2019), through the use of Jupyter Notebooks (Perez *et al.* 2015). This allows researchers to readily use our framework and also to contribute to the development of inversion codes that incorporate geophysical, geological and petrophysical data to yield meaningful solutions.

## ACKNOWLEDGEMENTS

We thank Tim Munday (CSIRO) for making the Bookpurnong data sets available. We also thank all of the developers who have contributed to the SimPEG package (<https://simpeg.xyz>). In particular we thank Lindsey J. Heagy for her development of architectural features that allowed us to use flexible regularization functions and also to connect forward modellings that were based upon different physics, and Seogi Kang and Dominique Fournier for their implementation of the parallelized stitched EM1D code in SimPEG as well as for the many fruitful conversations. A special thanks to Rowan Cockett as a main founder of the SimPEG project, and to the entire SimPEG group for their support and for many fruitful conversations.

## REFERENCES

- Ascher, U. & Greif, C., 2011. *A First Course in Numerical Methods*, Society for Industrial and Applied Mathematics.
- Astic, T., 2019. Open Source Software for Petrophysically and Geologically Guided Geophysical Inversion, Available at: <https://doi.org/10.5281/zenodo.2564169>, Accessed: 2019-02-14.
- Astic, T. & Chouteau, M., 2019. Geological interpretation of the northern flank of the Matagami camp, Québec, using gravity and magnetic inversions, *Can. J. Earth Sci.*, **56**(5), 471–482.
- Astic, T. & Oldenburg, D.W., 2018. Petrophysically guided geophysical inversion using a dynamic gaussian mixture model prior, in *SEG Technical Program Expanded Abstracts 2018*, pp. 2312–2316.
- Bosch, M., Carvajal, C., Rodrigues, J., Torres, A., Aldana, M. & Sierra, J., 2009. Petrophysical seismic inversion conditioned to well-log data: methods and application to a gas reservoir, *Geophysics*, **74**(2), O1–O15.
- Carter-McAuslan, A., Lelièvre, P.G. & Farquharson, C.G., 2015. A study of fuzzy c-means coupling for joint inversion, using seismic tomography and gravity data test scenarios, *Geophysics*, **80**(1), W1–W15.
- Cockett, R., Kang, S., Heagy, L.J., Pidlisecky, A. & Oldenburg, D.W., 2015. SimPEG: an open source framework for simulation and gradient based parameter estimation in geophysical applications, *Comput. Geosci.*, **85**, 142–154.
- Dempster, A.P., Laird, N.M. & Rubin, D.B., 1977. Maximum likelihood from incomplete data via the {EM} algorithm, *J. R. Stat. Soc., B*, **39**(1), 1–38.
- Farquharson, C.G. & Oldenburg, D.W., 2004. A comparison of automatic techniques for estimating the regularization parameter in non-linear inverse problems, *J. geophys. Int.*, **156**(3), 411–425.
- Fournier, D., 2015. A cooperative magnetic inversion method with Lp-norm regularization, *Master's thesis*, University of British Columbia.
- Fournier, D. & Oldenburg, D.W., 2019. Inversion using spatially variable mixed  $\ell_p$  norms, *J. geophys. Int.*, **218**(1), 268–282.
- Fullagar, P.K., Pears, G.A. & McMonnies, B., 2008. Constrained inversion of geologic surfaces pushing the boundaries, *The Leading Edge*, **27**(1), 98–105.
- Giraud, J., Pakyuz-Charrier, E., Jessell, M., Lindsay, M., Martin, R. & Ogarro, V., 2017. Uncertainty reduction in joint inversion using geologically conditioned petrophysical constraints, *Geophysics*, **82**(6), 1–61.

- Grana, D. & Della Rossa, E., 2010. Probabilistic petrophysical-properties estimation integrating statistical rock physics with seismic inversion, *Geophysics*, **75**(3), O21.
- Grana, D., Fjeldstad, T. & Omre, H., 2017. Bayesian Gaussian mixture linear inversion for geophysical inverse problems, *Math. Geosci.*, **49**(4), 493–515.
- Granek, J., 2011. Computing geologically consistent models from geophysical data, *Master's thesis*, University of British Columbia.
- Haber, E., 2014. *Computational Methods in Geophysical Electromagnetics*, Vol. 1, SIAM.
- Heagy, L.J., Cockett, R., Kang, S., Rosenkjaer, G.K. & Oldenburg, D.W., 2017. A framework for simulation and inversion in electromagnetics, *Comput. Geosci.*, **107**, 1–19.
- Holland, K., Jolly, I., McEwan, K., White, M., Berens, V., Souter, N. & Osmond, G., 2008. The 'bookpurnong experiment': will groundwater management and flooding improve the health of the floodplain vegetation?, in *Proceedings of the 2nd International Salinity Forum*, 31 March–3 April 2008.
- Kang, S., Cockett, R., Heagy, L.J. & Oldenburg, D.W., 2015. Moving between dimensions in electromagnetic inversions, in *SEG Technical Program Expanded Abstracts 2015*, pp. 5000–5004.
- Kang, S., Heagy, L.J., Cockett, R. & Oldenburg, D.W., 2017. Exploring nonlinear inversions: a 1D magnetotelluric example, *The Leading Edge*, **36**(8), 696–699.
- Kang, S., Heagy, L.J., Cockett, R. & Oldenburg, D.W., 2018. SimPEG-EM1D: gradient-based 1D inversion software for large-scale airborne electromagnetic data, in *Proceedings of the 2018 AGU Fall Meeting*, American Geophysical Union.
- Lelièvre, P.G., Oldenburg, D.W. & Williams, N.C., 2009. Integrating geological and geophysical data through advanced constrained inversions, *Explor. Geophys.*, **40**(4), 334–341.
- Lelièvre, P.G., Farquharson, C.G. & Hurich, C.A., 2012. Joint inversion of seismic traveltimes and gravity data on unstructured grids with application to mineral exploration, *Geophysics*, **77**(1), K1–K15.
- Li, Y. & Oldenburg, D.W., 2000. Incorporating geological dip information into geophysical inversions, *Geophysics*, **65**(1), 148–157.
- Li, Y., Melo, A., Martinez, C. & Sun, J., 2019. Geology differentiation: A new frontier in quantitative geophysical interpretation in mineral exploration, *The Leading Edge*, **38**(1), 60–66.
- Linde, N., Renard, P., Mukerji, T. & Caers, J., 2015. Geological realism in hydrogeological and geophysical inverse modeling: a review, *Adv. Water Resour.*, **86**, 86–101.
- Lochbühler, T., Vrugt, J.A., Sadegh, M. & Linde, N., 2015. Summary statistics from training images as prior information in probabilistic inversion, *J. geophys. Int.*, **201**(1), 157–171.
- McMillan, M.S., Schwarzbach, C., Haber, E. & Oldenburg, D.W., 2015. 3D parametric hybrid inversion of time-domain airborne electromagnetic data, *Geophysics*, **80**(6), K25–K36.
- Mehanee, S., Golubev, N. & Zhdanov, M.S., 2005. Weighted Regularized Inversion of Magnetotelluric Data, in *SEG Technical Program Expanded Abstracts 1998*, pp. 481–484.
- Moorkamp, M., Lelièvre, P.G., Linde, N. & Khan, A., 2016. *Integrated Imaging of the Earth: Theory and Applications*, Vol. **218**, John Wiley & Sons.
- Murphy, K.P., 2012. *Machine Learning: A Probabilistic Perspective*, The MIT Press.
- Nocedal, J. & Wright, S.J., 2006. *Numerical Optimization*, 2nd edn, Springer.
- Oldenburg, D.W. & Li, Y., 1999. Estimating depth of investigation in dc resistivity and ip surveys, *Geophysics*, **64**(2), 403–416.
- Oldenburg, D.W. & Li, Y., 2005. Inversion for applied geophysics: a tutorial, in *Near-Surface Geophysics*, ed. Butler, D.K., pp. 89–150, Society of Exploration Geophysicists.
- Paasche, H. & Troncicke, J., 2007. Cooperative inversion of 2d geophysical data sets: a zonal approach based on fuzzy c-means cluster analysis, *Geophysics*, **72**(3), A35–A39.
- Parker, R.L., 1977. Understanding inverse theory, *Ann. Rev. Earth planet. Sci.*, **5**(1), 35–64.
- Pedregosa, F. et al., 2011. Scikit-learn: machine learning in Python, *J. Mach. Learn. Res.*, **12**, 2825–2830.
- Perez, F., Berkeley, L. & Granger, B.E., 2015. *Project Jupyter: Computational Narratives as the Engine of Collaborative Data Science*, pp. 1–24.
- Raiffa, H. & Schlaifer, R., 1961. Applied statistical decision theory, Studies in managerial economics, Division of Research, Graduate School of Business Administration, Harvard University.
- Sun, J. & Li, Y., 2015. Multidomain petrophysically constrained inversion and geology differentiation using guided fuzzy c-means clustering, *Geophysics*, **80**(4), ID1.
- Sun, J. & Li, Y., 2016. Joint inversion of multiple geophysical data using guided fuzzy c-means clustering, *Geophysics*, **81**(3), ID37–ID57.
- Sun, J. & Li, Y., 2017. Joint inversion of multiple geophysical and petrophysical data using generalized fuzzy clustering algorithms, *J. geophys. Int.*, **208**(2), 1201–1216.
- Sun, J. & Li, Y., 2018. Magnetization clustering inversion – part 1: building an automated numerical optimization algorithm, *Geophysics*, **83**(5), J61–J73.
- Tarantola, A., 2005. *Inverse Problem Theory and Methods for Model Parameter Estimation*, Vol. **89**, Society for Industrial and Applied Mathematics.
- Tikhonov, A.N. & Arsenin, V.Y., 1977. *Solutions of Ill-Posed Problems*, John Wiley & Sons, Translated from the Russian, Preface by translation editor Fritz John, Scripta Series in Mathematics.
- Viezzoli, A., Christiansen, A.V., Auken, E. & Sørensen, K., 2008. Quasi-3D modeling of airborne tem data by spatially constrained inversion, *Geophysics*, **73**(3), F105–F113.
- Viezzoli, A., Auken, E. & Munday, T., 2009. Spatially constrained inversion for quasi 3D modelling of airborne electromagnetic data – an application for environmental assessment in the lower murray region of south australia, *Explor. Geophys.*, **40**(2), 173–183.
- Viezzoli, A., Munday, T., Auken, E. & Christiansen, A.V., 2010. Accurate quasi 3D versus practical full 3D inversion of AEM data – the Bookpurnong case study, *Preview*, **2010**(149), 23–31.
- Ward, S.H. & Hohmann, G.W., 1988. *Electromagnetic Theory for Geophysical Applications*, SEG Library, pp. 130–311.
- Williams, N.C., 2008. Geologically-constrained UBC-GIF gravity and magnetic inversions with examples from the Agnew-Wiluna greenstone belt, PhD thesis, University of British Columbia, Western Australia.
- Wu, X., 2017. Structure-, stratigraphy- and fault-guided regularization in geophysical inversion, *J. geophys. Int.*, **210**(1), 184–195.
- Yang, D., 2017. Airborne electromagnetic data inversion for floodplain salinization investigation at Bookpurnong, South Australia, [https://figshare.com/articles/Airborne\\_electromagnetic\\_data\\_inversion\\_for\\_floodplain\\_salinization\\_investigation\\_at\\_Bookpurnong\\_South\\_Australia/5071573](https://figshare.com/articles/Airborne_electromagnetic_data_inversion_for_floodplain_salinization_investigation_at_Bookpurnong_South_Australia/5071573), Accessed: 2019-02-01.
- Zhdanov, M.S. & Lin, W., 2017. Adaptive multinary inversion of gravity and gravity gradiometry data, *Geophysics*, **82**(6), G101–G114.

## APPENDIX A: FROM A POSTERIOR WITH GAUSSIAN DISTRIBUTIONS TO A LEAST-SQUARES OBJECTIVE-FUNCTION

In this appendix, we develop the operations necessary to go from a posterior distribution formulation to an objective function in the case of the Tikhonov inverse problem.

We stated in Section 2 that a least-squares term corresponds to a Gaussian prior. For a vector of parameters, the multivariate Gaussian distribution, denoted by  $\mathcal{N}$ , with mean  $\mu$  and covariance  $\Sigma$  (denoted by  $\sigma^2$  for a unidimensional Gaussian, with  $\sigma$  being the standard-deviation) is:

$$\mathcal{N}(\mathbf{m}|\mu, \Sigma) = \frac{\exp(-\frac{1}{2} \|\Sigma^{-\frac{1}{2}}(\mathbf{m} - \mu)\|_2^2)}{\sqrt{(2\pi)^n \det(\Sigma)}}. \quad (\text{A1})$$

Going from a posterior distribution to an objective function is done by taking the negative natural logarithm of the posterior. Taking the negative logarithm of a single Gaussian term gives:

$$-\log(\mathcal{N}(\mathbf{m}|\boldsymbol{\mu}, \Sigma)) = \frac{1}{2} \|\Sigma^{-\frac{1}{2}}(\mathbf{m} - \boldsymbol{\mu})\|_2^2 + \log(\sqrt{(2\pi)^n \det(\Sigma)}). \quad (\text{A2})$$

The first term obtained in eq. (A2) is recognizable as a least-squares misfit. The second term does not depend on  $\mathbf{m}$  and is thus a constant. As constants do not play any role in the optimization, it is thus discarded in the objective function.

For the full Tikhonov geophysical objective function  $\mathcal{P}(\mathbf{m}|\mathbf{d}_{\text{obs}})$ , the full posterior distributions can be written, using the Bayes rules, as:

$$\mathcal{P}(\mathbf{m}|\mathbf{d}_{\text{obs}}) = \frac{\mathcal{P}(\mathbf{d}_{\text{obs}}|\mathbf{m})\mathcal{P}(\mathbf{m})}{\mathcal{P}(\mathbf{d}_{\text{obs}})} \quad (\text{A3})$$

with:

$$\mathcal{P}(\mathbf{d}_{\text{obs}}|\mathbf{m}) = \mathcal{N}(\mathcal{F}[\mathbf{m}]|\mathbf{d}_{\text{obs}}, (W_d^T W_d)^{-1}) \quad (\text{A4})$$

$$\mathcal{P}(\mathbf{m}) = \mathcal{N}(\mathbf{m}|\mathbf{m}_{\text{ref}}, (\beta\alpha_s W_s^T W_s)^{-1})\mathcal{N}(\mathbf{m}|\mathbf{m}_{\text{ref}}, (\beta L^T L)^{-1}), \quad (\text{A5})$$

where the finite difference operator  $L$  summarizes the first or second derivatives in all directions with their respective  $\{\alpha\}$  scale. We, respectively, express the data misfit, the smallness and smoothness probability functions as Gaussian distributions:

$$\mathcal{N}(\mathbf{d}_{\text{obs}}|\mathbf{m}, (W_d^T W_d)^{-1}) = \frac{\exp(-\frac{1}{2}\|\mathcal{F}[\mathbf{m}] - \mathbf{d}_{\text{obs}}\|_2^2)}{\sqrt{(2\pi)^n \det([W_d^T W_d]^{-1})}} \quad (\text{A6})$$

$$\mathcal{N}(\mathbf{m}|\mathbf{m}_{\text{ref}}, (\beta\alpha_s W_s^T W_s)^{-1}) = \frac{\exp(-\frac{1}{2}\|\mathbf{m} - \mathbf{m}_{\text{ref}}\|_2^2)^{\beta\alpha_s}}{\sqrt{(2\pi)^n \det([\beta\alpha_s W_s^T W_s]^{-1})}} \quad (\text{A7})$$

$$\mathcal{N}(\mathbf{m}|\mathbf{m}_{\text{ref}}, (\beta L^T L)^{-1}) = \frac{\exp(-\frac{1}{2}\|L(\mathbf{m} - \mathbf{m}_{\text{ref}})\|_2^2)^{\beta}}{\sqrt{(2\pi)^n \det([\beta L^T L]^{-1})}}. \quad (\text{A8})$$

The objective function is obtained by applying the negative natural logarithm to the posterior distribution described in eq. (A3). The summation form is simply a consequence of the fundamental property of the logarithm function, the multiplication becomes an addition in this new space. It follows:

$$\Phi(\mathbf{m}) = \frac{1}{2} \|\mathcal{F}[\mathbf{m}] - \mathbf{d}_{\text{obs}}\|_2^2 + \frac{\beta\alpha_s}{2} \|\mathbf{m} - \mathbf{m}_{\text{ref}}\|_2^2 + \frac{\beta}{2} \|L(\mathbf{m} - \mathbf{m}_{\text{ref}})\|_2^2 + \text{Constant}. \quad (\text{A9})$$

The constant term in eq. (A9) contains the constant terms for each Gaussian distribution and the constant term  $\log(\mathcal{P}(\mathbf{d}_{\text{obs}}))$ .

This completes the detailed operation to go from a posterior distribution to an objective function formulation of the Tikhonov inverse problem.

## APPENDIX B: CONJUGATE PRIOR VERSUS SEMI-CONJUGATE PRIOR FOR MEANS AND VARIANCES PRIOR INFORMATION FOR A SINGLE UNIVARIATE GAUSSIAN DISTRIBUTION

In this section, we elaborate on the difference between a full and a semi-conjugate prior approach for estimating the parameters of a

single, 1-D, Gaussian distribution. After giving the needed definitions for the full conjugate prior and how the MAP-EM updates are affected, we illustrate the difference between semi and full conjugate through the example displayed in Fig. B1.

We first need some definitions. The full conjugate prior for the mean and variance follows a normal-inverse-gamma distribution (Murphy 2012). As in Section 3.2, we use the normal-inverse chi-squared reparametrization:

$$\mathcal{P}(\mu, \sigma^2) = \mathcal{N}\mathcal{X}^{-2}(\mu, \sigma^2|\mu_{\text{prior}}, K, N, \sigma_{\text{prior}}^2), \quad (\text{B1})$$

with the confidences  $K$  in the prior mean and  $N$  in the prior variance:

$$K = (\kappa\pi_{\text{prior}} V)^{-1} \quad (\text{B2})$$

$$N = v\pi_{\text{prior}} V. \quad (\text{B3})$$

The update of the means in MAP-EM algorithm stays the same as for the semi-conjugate prior approach, as detailed in eq. (32). On the contrary the update of the variances now requires an additional term  $s_{\bar{\mu}_j}^{(k)}$  to account for the difference between the observed and the prior means:

$$\sigma_j^{(k)} = \frac{V_j^{(k)} \sigma_{\bar{\mu}_j}^{(k)} + v_j \pi_j \text{prior} V \sigma_{j \text{ prior}}^2 + s_{\bar{\mu}_j}^{(k)}}{V_j^{(k)} + v_j \pi_j \text{prior} V}, \quad (\text{B4})$$

with:

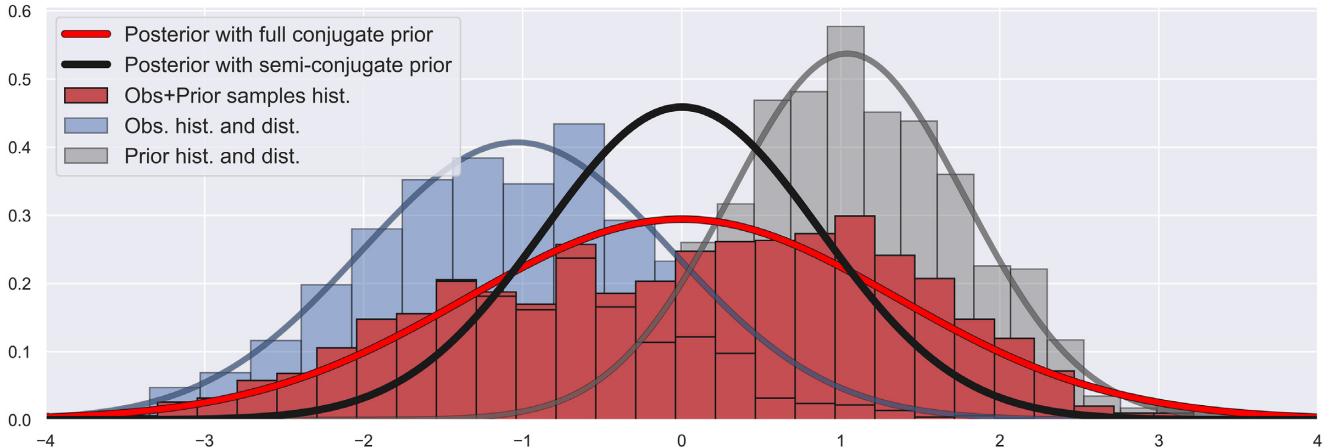
$$\sigma_{\bar{\mu}_j}^{(k)} = \frac{1}{V_j^{(k)}} \sum_{i=1}^n v_i n_{ij}^{(k)} (m_i - \bar{m}_j^{(k)})^2 \quad (\text{B5})$$

$$s_{\bar{\mu}_j}^{(k)} = \frac{\kappa_j V_j^{(k)}}{\kappa_j + V_j^{(k)}} (\bar{m}_j^{(k)} - \mu_{j \text{ prior}})^2. \quad (\text{B6})$$

We illustrate those concepts with Fig. B1. Let us consider the following setup. We have a set of observations (in our framework, this was our geophysical model), represented through their histogram in blue. The Gaussian distribution in blue represents the MLE parameters (estimated without any prior information). We now add prior information in the form of a Gaussian distribution, in grey in the Fig. B1. We set the confidence parameters  $\kappa$  and  $v$  to unity, meaning that we have equal confidence in the observations and the prior ( $\zeta$  is irrelevant here as we have only one Gaussian distribution). This equal confidence can be represented by having an equal number of samples from the prior than in the observed set. The histogram of this prior synthetic samples set is shown in grey.

The full conjugate prior approach can be understood as fitting a Gaussian distribution on the data set formed by merging the observed and synthetic observations; this is represented in Fig. B1 in red. The full conjugate MAP distribution, also in red, is well centred between the two observed and prior distribution as expected, and so is the semi-conjugate distribution (in black). However the variance of the red histogram, and thus of the posterior distribution with full conjugate prior, is higher than the variance of either the observed or prior distribution. This is due to the difference in the means of the two distributions, which the full conjugate prior approach accounts for (see eqs B4 and B6). The semi-conjugate prior approach considers the means and variances independently (see eqs 30 and 31). The MAP mean estimates are the same for both priors.

The semi-conjugate prior approach seems a better choice in the context of geophysical inversion. As our goal with this framework is to differentiate various geological units, by guiding the geophysical model to reproduce certain features, the full conjugate prior



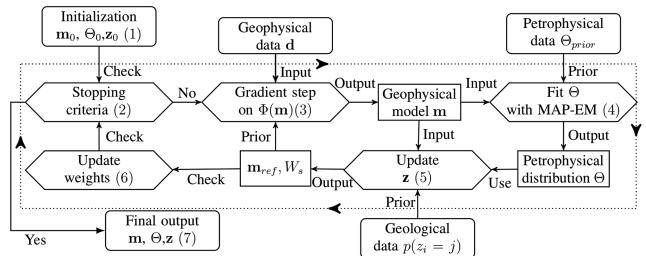
**Figure B1.** A comparison of the MLE estimations of the mean and variance for a single 1-D Gaussian distribution (no prior), semi-conjugate prior and full conjugate prior MAPs. We have observed samples in blue, along with their estimated MLE Gaussian distribution, and a prior distribution in grey. Using confidence values of unity in the prior is similar to having an equal number of samples from the prior and in the observed data sets. A synthetic prior sample set is represented as the grey histogram. The posterior distribution with a semi-conjugate prior is in black and is seen as an average of the parameters of the observed and prior distributions. The posterior distribution with a full conjugate prior is in red. The red histogram is obtained by merging the observed samples with the synthetic prior samples. This histogram corresponds to the posterior distribution with a full conjugate prior. Note that the variance is larger than for either of the two original distributions; this results because of the difference in the means. Hist. stands for histogram; Dist. stands for distribution; Obs. stands for observed samples.

approach is sometimes detrimental as it can drive the variance further from the prior than what is currently seen in the model at each iteration.

## APPENDIX C: PSEUDOCODE ALGORITHM

We give in algorithm 1 the details of our implementation of this framework. The optimization notions of inexact Gauss–Newton and backtracking line search are detailed in Ascher & Greif (2011), or in Haber (2014) for their geophysical applications.

Fig. C1 graphically summarizes in a flowchart how our framework loops over the various data sets to produce a final geophysical model with the desired petrophysical distribution and geological features.



**Figure C1.** A visual pseudocode of our iterative cyclic framework to include geological and petrophysical information into geophysical inversion. The numbers correspond to the steps displayed in algorithm 1.

---

**Algorithm 1** This algorithm finds a geophysical model reproducing the geophysical, petrophysical and geological data.

---

1 **Initialization:**

- Input:
  - Initial geophysical model  $\mathbf{m}^{(0)}$ , petrophysical distributions  $\Theta^{(0)}$  and geological model  $\mathbf{z}^{(0)}$ .

• Parameters:

- *Objective function*: data's noise matrix  $W_d$ , trade-off parameter  $\beta^{(0)}$ , prior matrices and weights  $\{\alpha\}$ ,  $\{W\}$ .
- *Localized prior*: specific  $P(z_i)$  for available locations  $i \in \{1..n\}$ , local weights  $\{\mathbf{w}_i\}_{i=1..n}$ .
- *Confidences in the petrophysical prior*:  $\{\kappa_j, v_j, \zeta_j\}_{j=1..c}$  for the means, variances and global proportions.
- *Optimization*:  $\beta$ -cooling factor  $\gamma (> 1)$ , sufficient decrease rate  $\tau (\leq 1)$ .

• Output:

- $\mathbf{m}$ ,  $\Theta$ ,  $\mathbf{z}$ .

2 **while**  $\Phi_d > \Phi_d^*$  and  $\Phi_s > \Phi_s^*$  **do**3   **Objective-Function Descent Step:**

- Compute a model perturbation  $\delta\mathbf{m}$  for 1 with the smallness defined in 20 using an inexact Gauss-Newton approach.
- Backtracking line search with a Wolfe condition to find a step-size  $\eta$  that satisfy a sufficient decrease of the objective function.
- Return  $\mathbf{m}^{(t)} = \mathbf{m}^{(t-1)} + \eta\delta\mathbf{m}$ .

4   **Update Petrophysical Distribution**

- Fit a new GMM  $\Theta^{(t)}$  on  $\mathbf{m}^{(t)}$  with a confidence in the prior determined by the hyper-parameters  $\{\zeta\}$ ,  $\{\kappa\}$ ,  $\{v\}$  such as in eqs 27, 32 and 34 until no sufficient increase of 24 is observed.

5   **Classification:**

- Compute the membership  $\mathbf{z}^{(t)}$  of the current model  $\mathbf{m}^{(t)}$  as in eq. 21 using  $\Theta^{(t)}$  as the classifier.
- Update  $\mathbf{m}_{\text{ref}}$  and  $W_s$  according to 22 and 23 respectively using  $\mathbf{z}^{(t)}$ .

6   **Update weights:**

```
if  $\Phi_d^{(t)} > \Phi_d^*$  and  $\Phi_s^{(t)} > \Phi_s^*$  then
```

- Decrease  $\beta$ :  $\beta^{(t)} = \frac{\beta^{(t-1)}}{\gamma}$

```
else if  $\Phi_d^{(t)} \leq \Phi_d^*$  and  $\Phi_s > \Phi_s^*$  then
```

- Increase  $\alpha_s$ :  $\alpha_s^{(t)} = \alpha_s^{(t-1)} \times \frac{\Phi_d^*}{\Phi_d^{(t)}}$

```
if (optional)  $\Phi_d^{(t)} \leq \Phi_d^*$  and  $\Phi_s > \Phi_s^*$  and  $\mathbf{z}^{(t)} == \mathbf{z}^{(t-1)}$  then
```

- Include  $\mathbf{m}_{\text{ref}}$  in Smoothness

7 **end**