

A VARIATIONAL BAYES APPROACH TO ADAPTIVE CHANNEL-GAIN CARTOGRAPHY

Donghoon Lee and Georgios B. Giannakis

Dept. of ECE and Digital Technology Center, University of Minnesota, USA
Emails: {leex6962, georgios}@umn.edu

ABSTRACT

Channel-gain cartography relies on sensor measurements to construct maps providing the attenuation profile between arbitrary transmitter-receiver locations. State-of-the-art on this subject includes tomography-based approaches, where shadowing effects are modeled by the weighted integral of a spatial loss field (SLF) that captures the propagation environment. To learn SLFs exhibiting statistical heterogeneity induced by spatially diverse propagation environments, the present work develops a Bayesian approach comprising a piecewise homogeneous SLF with an underlying hidden Markov random field model. Built on a variational Bayes scheme, the novel approach yields efficient field estimators at affordable complexity. In addition, a data-adaptive sensor selection algorithm is developed to collect informative measurements for effective learning of the SLF. Numerical tests demonstrate the capabilities of the novel approach.

Index Terms— channel-gain cartography, radio tomography, variational Bayes, active learning

1. INTRODUCTION

Based on measurements collected by a network of spatially distributed sensors, channel-gain (CG) cartography constructs maps providing channel-state information for links even between locations where no sensors are present [14]. Such maps can be employed by cognitive radio communication networks to monitor and control the interference that the secondary network inflicts to primary users that do not transmit – a setup encountered with television broadcast systems [2, 26, 6, 13]. The non-collaborative nature of these primary users precludes any direct form of pilot-based or blind channel estimation between secondary transmitters and primary receivers.

Existing methods for channel-gain cartography build upon the intuitive principle that spatially close radio links exhibit similar shadowing [1]. Most of these methods adopt a tomographic approach [22], where shadowing is modeled as the weighted integral of an unknown spatial loss field (SLF) capturing the absorption induced by objects located across the propagation medium [22, 25, 5, 17]. The weights

in the integral are determined by a function depending on transmitter-receiver locations that is either selected based on heuristic criteria [22, 9], or blindly learned using non-parametric kernel regression [23]. A channel gain map can thus be obtained once the SLF has been estimated.

Conventionally, the SLF is learned via regularized least-squares (LS) methods tailored to the propagation environment [17, 9, 25]. However, these approaches are less effective when the propagation environment is spatially heterogeneous due to a combination of free space and objects present in different sizes and materials (the typical setup in e.g., an urban area), which subsequently induces statistical diversity in the SLF. To deal with such heterogeneous environments, we proposed in [16] a Bayesian approach to learn piecewise homogeneous SLFs through a hidden Markov random field (MRF) model [11] obtained via Markov chain Monte Carlo (MCMC) [7]. However, MCMC can be computationally demanding and limits its use to a small-scale problem.

To provide efficient field estimators with low computational complexity, we derive here a variational Bayes estimator to approximate the analytically intractable minimum mean-square error (MMSE) or maximum a posteriori (MAP) estimators. We further develop a data-adaptive sensor selection method to permeate benefits of active learning [18], and collect informative data that lower SLF uncertainty.

Notation. Matrix $\mathbf{I}_n$ is the $n \times n$ identity matrix; superscript $\top$ stands for transposition; and $|\cdot|$ for set cardinality.

2. MODEL AND PROBLEM STATEMENT

Consider a set of sensors deployed over a two-dimensional geographical area indexed by a set $\mathcal{A} \subset \mathbb{R}^2$. After averaging out small-scale fading effects, the channel-gain (CG) over a link between a transmitter located at $\mathbf{x} \in \mathcal{A}$, and a receiver located at $\mathbf{x}' \in \mathcal{A}$, can be represented (in dB) as

$$g(\mathbf{x}, \mathbf{x}') = g_0 - \gamma 10 \log_{10} d(\mathbf{x}, \mathbf{x}') - s(\mathbf{x}, \mathbf{x}') \quad (1)$$

where g_0 is the path gain at unit distance; $d(\mathbf{x}, \mathbf{x}') := \|\mathbf{x} - \mathbf{x}'\|_2$ is the distance between the transceivers at $\mathbf{x}$ and $\mathbf{x}'$; γ is the pathloss exponent; and $s(\mathbf{x}, \mathbf{x}')$ denotes the attenuation due to shadow fading. A tomographic shadow fading model typically adopted in CG cartography is [22, 9, 17]

The work in this paper was supported by NSF grants 1442686, 1508993, 1509040.

$$s(\mathbf{x}, \mathbf{x}') \simeq \sum_{i=1}^{N_g} w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}_i) f(\tilde{\mathbf{x}}_i). \quad (2)$$

where $\{\tilde{\mathbf{x}}_i\}_{i=1}^{N_g}$ is a grid of points over $\mathcal{A}$, $f : \mathcal{A} \rightarrow \mathbb{R}$ denotes the *spatial loss field* (SLF) capturing the attenuation at each location, and $w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}})$ is the weight function modeling the influence of the SLF at $\tilde{\mathbf{x}}$ to the shadowing experienced by the link $\mathbf{x}-\mathbf{x}'$. Examples of the weight function include the *normalized ellipse model* taking the form [25]

$$w(\mathbf{x}, \mathbf{x}', \tilde{\mathbf{x}}) := \begin{cases} 1/\sqrt{d(\mathbf{x}, \mathbf{x}')}, & \text{if } d(\mathbf{x}, \tilde{\mathbf{x}}) + d(\mathbf{x}', \tilde{\mathbf{x}}) \\ & < d(\mathbf{x}, \mathbf{x}') + \lambda \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

where $\lambda > 0$ is a tunable parameter. The value of λ is commonly set to half the wavelength to assign non-zero weights only within the first Fresnel zone. Overall, the model in (2) shows how nature and the spatial distribution of obstructions in the propagation medium influence the attenuation between a pair of locations.

To estimate the CG map, N sensors located at $\{\mathbf{x}_1 \cdots \mathbf{x}_N\} \in \mathcal{A}$ collaboratively obtain CG measurements. At time slot τ , the radios indexed by $n(\tau)$ and $n'(\tau)$ measure the channel-gain $g_\tau := g(\mathbf{x}_{n(\tau)}, \mathbf{x}_{n'(\tau)})$ by exchanging pilot sequences, where $n(\tau), n'(\tau) \in \{1, \dots, N\}$. It is supposed that g_0 and γ have been estimated during a calibration phase. After subtracting these from g_τ , the shadowing estimate $\tilde{s}_\tau := \tilde{s}(\mathbf{x}_{n(\tau)}, \mathbf{x}_{n'(\tau)}) := s(\mathbf{x}_{n(\tau)}, \mathbf{x}_{n'(\tau)}) + \nu_\tau$ is obtained, where ν_τ denotes estimation and measurement noise.

Given $\tilde{\mathbf{s}}_t := [\tilde{s}_1, \dots, \tilde{s}_t]^\top \in \mathbb{R}^t$, along with the known set of link locations $\{(\mathbf{x}_{n(\tau)}, \mathbf{x}_{n'(\tau)})\}_{\tau=1}^t$ and the weight function w , the goal is to estimate $g(\mathbf{x}, \mathbf{x}')$ between any pair of locations $(\mathbf{x}, \mathbf{x}') \in \mathcal{A}$. To this end, it suffices to estimate f , or equivalently $\mathbf{f} := [f(\tilde{\mathbf{x}}_1), \dots, f(\tilde{\mathbf{x}}_{N_g})]^\top \in \mathbb{R}^{N_g}$. Afterwards, the arbitrary channel-gain $g(\mathbf{x}, \mathbf{x}')$ can be obtained by substituting (2) into (1) and replacing f with its estimate.

3. ADAPTIVE BAYESIAN CG CARTOGRAPHY

In this section, we propose a variational Bayes approach for inference with a two-layer Bayesian SLF model; and a data-adaptive sensor selection method via uncertainty sampling.

3.1. Field estimation via variational Bayes

Let $\mathcal{A}$ consist of two disjoint homogeneous regions $\mathcal{A}_0 := \{\mathbf{x} | \mathbb{E}[f(\mathbf{x})] = \mu_{f_0}, \text{Var}[f(\mathbf{x})] = \sigma_{f_0}^2, \mathbf{x} \in \mathcal{A}\}$ and $\mathcal{A}_1 := \{\mathbf{x} | \mathbb{E}[f(\mathbf{x})] = \mu_{f_1}, \text{Var}[f(\mathbf{x})] = \sigma_{f_1}^2, \mathbf{x} \in \mathcal{A}\}$, giving rise to a hidden label field $\mathbf{z} := [z(\tilde{\mathbf{x}}_1), \dots, z(\tilde{\mathbf{x}}_{N_g})]^\top \in \{0, 1\}^{N_g}$ of binary labels with $z(\tilde{\mathbf{x}}_i) = k$ if $\tilde{\mathbf{x}}_i \in \mathcal{A}_k \forall i$, and $k = 0, 1$. We then model the conditional distribution of $f(\tilde{\mathbf{x}}_i)$ as

$$p(f(\tilde{\mathbf{x}}_i) | z(\tilde{\mathbf{x}}_i) = k) = \mathcal{N}(\mu_{f_k}, \sigma_{f_k}^2). \quad (4)$$

and let $\mathbf{z}$ adhere to an Ising model [24], which is a binary version of the discrete MRF with the Potts model [11]. This

Ising model captures the dependency among spatially correlated labels, and by the Hammersley-Clifford theorem [10], it corresponds to drawing labels from a Gibbs distribution

$$p(\mathbf{z}; \beta) = \frac{1}{C(\beta)} \exp \left[\beta \sum_{i=1}^{N_g} \sum_{j \in \mathcal{N}(\tilde{\mathbf{x}}_i)} \delta(z(\tilde{\mathbf{x}}_j) - z(\tilde{\mathbf{x}}_i)) \right] \quad (5)$$

where $\mathcal{N}(\tilde{\mathbf{x}}_i)$ is a set of indices associated with one-hop neighbors of $\tilde{\mathbf{x}}_i$ on the rectangular grid, β is the granularity coefficient to control the degree of homogeneity in $\mathbf{z}$, $\delta(\cdot)$ is the Kronecker delta function, and $C(\beta) := \sum_{\mathbf{z} \in \mathcal{Z}} \exp \left[\beta \sum_{i=1}^{N_g} \sum_{j \in \mathcal{N}(\tilde{\mathbf{x}}_i)} \delta(z(\tilde{\mathbf{x}}_j) - z(\tilde{\mathbf{x}}_i)) \right]$ is the partition function with $\mathcal{Z} := \{0, 1\}^{N_g}$. By assuming conditional independence of $\{f(\tilde{\mathbf{x}}_i)\}_{i=1}^{N_g}$ given $\mathbf{z}$, the resulting model is referred to as the mixture of independent Gaussians (MIG) having a Potts prior model [3] with two labels.

Suppose ν_τ is independent and identically distributed (i.i.d) Gaussian with zero mean and variance σ_ν^2 , and let $\boldsymbol{\theta}$ denote the known parameter vector including σ_ν^2 , β , and $\boldsymbol{\theta}_f := [\mu_{f_0}, \mu_{f_1}, \sigma_{f_0}^2, \sigma_{f_1}^2]^\top$ to ease exposition. The weight matrix $\mathbf{W}_t \in \mathbb{R}^{N_g \times t}$ is constructed with columns equal to $\mathbf{w}_\tau^{(n, n')} := [w(\mathbf{x}_{n(\tau)}, \mathbf{x}_{n'(\tau)}, \tilde{\mathbf{x}}_1), \dots, w(\mathbf{x}_{n(\tau)}, \mathbf{x}_{n'(\tau)}, \tilde{\mathbf{x}}_{N_g})]^\top \in \mathbb{R}^{N_g}$ of the link $\mathbf{x}_{n(\tau)}-\mathbf{x}_{n'(\tau)}$ for $\tau = 1, \dots, t$. Then, one can cast Bayesian CG cartography through the joint posterior

$$p(\mathbf{f}, \mathbf{z} | \tilde{\mathbf{s}}_t; \boldsymbol{\theta}) \propto p(\tilde{\mathbf{s}}_t | \mathbf{f}; \sigma_\nu^2) p(\mathbf{f} | \mathbf{z}; \boldsymbol{\theta}_f) p(\mathbf{z}; \beta) \quad (6)$$

where $p(\tilde{\mathbf{s}}_t | \mathbf{f}; \sigma_\nu^2) \sim \mathcal{N}(\mathbf{W}_t^\top \mathbf{f}, \sigma_\nu^2 \mathbf{I}_t)$ is the data likelihood. By utilizing the posterior in (6), the MMSE estimator of $\mathbf{f}$ is $\hat{\mathbf{f}}_{\text{MMSE}} := \mathbb{E}[\mathbf{f} | \mathbf{z} = \hat{\mathbf{z}}_{\text{MAP}}, \tilde{\mathbf{s}}_t]$, where $\mathbf{z}$ is fixed to the marginal MAP estimate of $\mathbf{z}$, that is, $\hat{\mathbf{z}}_{\text{MAP}} = \arg \max_{\mathbf{z}} p(\mathbf{z} | \tilde{\mathbf{s}}_t)$.

Although the suggested estimators have been advocated in [12], analytical solutions are not available because the posterior (6) needed for marginalization or maximization has complex form. To bypass this challenge, we will consider approximate yet analytically tractable solutions from a surrogate distribution close to the posterior in (6). Variational Bayes (VB) is a family of techniques providing an analytical approximation to a complex distribution, that is referred to as *variational distribution*. A typical choice of an approximation criterion is to find the variational distribution minimizing the Kullback-Leibler (KL) divergence to a target distribution. The variational distribution is further assumed to belong to a certain family of distributions maintaining a simpler form of dependence between variables than the original one; see also [21] for the so-termed mean-field approximation.

Let q and $\mathcal{Q}$ denote the variational distribution and the associated family of distributions, respectively. Then, $q(\mathbf{f}, \mathbf{z})$ of the posterior in (6) can be found by solving $\min_{q(\mathbf{f}, \mathbf{z}) \in \mathcal{Q}} D_{\text{KL}}(q(\mathbf{f}, \mathbf{z}) \| p(\mathbf{f}, \mathbf{z} | \tilde{\mathbf{s}}_t; \boldsymbol{\theta}))$, where $D_{\text{KL}}(q \| p)$ is the KL divergence from q to p , or equivalently

$$(P1) \max_{q(\mathbf{f}, \mathbf{z}) \in \mathcal{Q}} \mathbb{E}_{q(\mathbf{f}, \mathbf{z})} [\ln p(\mathbf{f}, \mathbf{z}, \tilde{\mathbf{s}}_t; \boldsymbol{\theta})] - \mathbb{E}_{q(\mathbf{f}, \mathbf{z})} [\ln q(\mathbf{f}, \mathbf{z})]$$

since only $p(\mathbf{f}, \mathbf{z}, \tilde{\mathbf{s}}_t; \boldsymbol{\theta})$, which is equal to the right-hand side of (6), is available in practice, not the exact posterior $p(\mathbf{f}, \mathbf{z} | \tilde{\mathbf{s}}_t; \boldsymbol{\theta})$. To take into account the dependence between $\mathbf{f}$ and $\mathbf{z}$, the family $\mathcal{Q}$ is defined as

$$\mathcal{Q} := \left\{ q : q(\mathbf{f}, \mathbf{z}) := q(\mathbf{f} | \mathbf{z}) q(\mathbf{z}) = \prod_{i=1}^{N_g} q(f_i | z_i) \prod_{i=1}^{N_g} q(z_i) \right\} \quad (7)$$

with $f_i := f(\tilde{\mathbf{x}}_i)$ and $z_i := z(\tilde{\mathbf{x}}_i) \forall i$ for simplicity. Then, (P1) can be solved via coordinate descent w.r.t. each factor of $q(\mathbf{f}, \mathbf{z})$ in (7) [20]. The optimal solutions have the form

$$\ln q^*(f_i | z_i) = \mathbb{E}_{\prod_{j \neq i} q(f_j | z_j) q(\mathbf{z})} [\ln p(\mathbf{f}, \mathbf{z}, \tilde{\mathbf{s}}_t; \boldsymbol{\theta})] + c \forall i \quad (8)$$

$$\ln q^*(z_i) = \mathbb{E}_{q(\mathbf{f} | \mathbf{z}) \prod_{j \neq i} q(z_j)} [\ln p(\mathbf{f}, \mathbf{z}, \tilde{\mathbf{s}}_t; \boldsymbol{\theta})] + c \forall i \quad (9)$$

where c is a generic normalization constant. Note that the solutions in (8) and (9) are intertwined since the evaluation of the former requires the latter, and vice versa. For this reason, we can show that the optimal solutions can be obtained iteratively; that is $\forall i$, at iteration $\ell = 1, 2, \dots$, we have

$$q^{(\ell)}(f_i | z_i = k) = \mathcal{N}(\check{\mu}_{f_k}^{(\ell)}(\tilde{\mathbf{x}}_i), \check{\sigma}_{f_k}^2(\tilde{\mathbf{x}}_i)) \forall k \quad (10)$$

$$q^{(\ell)}(z_i = k) \propto \exp \left\{ -\frac{1}{2\sigma_{f_k}^2} \left[\check{\sigma}_{f_k}^2(\tilde{\mathbf{x}}_i) + \left(\check{\mu}_{f_k}^{(\ell)}(\tilde{\mathbf{x}}_i) \right)^2 - 2\mu_{f_k} \check{\mu}_{f_k}^{(\ell)}(\tilde{\mathbf{x}}_i) + \mu_{f_k}^2 \right] + \sum_{j \in \mathcal{N}(\tilde{\mathbf{x}}_i)} \beta q^{(\ell-1)}(z_j = k) \right\} \forall k \quad (11)$$

with

$$\check{\sigma}_{f_k}^2(\tilde{\mathbf{x}}_i) = \left[\frac{1}{\sigma_{f_k}^2} \sum_{\tau=1}^t [\mathbf{W}_t^\top]_{\tau, i}^2 + \frac{1}{\sigma_{f_k}^2} \right]^{-1} \quad (12)$$

$$\check{\mu}_{f_k}^{(\ell)}(\tilde{\mathbf{x}}_i) = \bar{f}_i^{(\ell-1)} + \check{\sigma}_{f_k}^2(\tilde{\mathbf{x}}_i) \times \left[\frac{\mu_{f_k} - \bar{f}_i^{(\ell-1)}}{\sigma_{f_k}^2} + \frac{1}{\sigma_{f_k}^2} \sum_{\tau=1}^t [\mathbf{W}_t^\top]_{\tau, i} (\check{s}_\tau - \bar{s}_\tau^{(\ell-1)}) \right], \quad (13)$$

where $\bar{f}_i^{(\ell)} := \sum_{k=0}^1 q^{(\ell)}(z_i = k) \check{\mu}_{f_k}^{(\ell)}(\tilde{\mathbf{x}}_i)$, and $\bar{s}_\tau^{(\ell)} := \sum_{i=1}^{N_g} [\mathbf{W}_t^\top]_{\tau, i} \bar{f}_i^{(\ell)}$.

Accordingly, the MAP estimator of $\mathbf{z}$ can be approximated as $\hat{z}_{\text{MAP}, i} = \arg \max_{z_i \in \{0, 1\}} q^*(z_i) \forall i$ and subsequently, the approximate MMSE estimator of $\mathbf{f}$ is given by $\hat{f}_{i, \text{MMSE}} = \mathbb{E}_{q^*(f_i | \hat{z}_{\text{MAP}, i})} [f_i] = \check{\mu}_{f_{\hat{z}_{\text{MAP}, i}}}^*(\tilde{\mathbf{x}}_i) \forall i$.

The VB algorithm to obtain $\hat{\mathbf{f}}_{\text{MMSE}} := \{\hat{f}_{i, \text{MMSE}}\}_{i=1}^{N_g}$, $q^*(\mathbf{f} | \mathbf{z})$, and $q^*(\mathbf{z})$ is tabulated as Alg. 1.

3.2. Adaptive data collection via uncertainty sampling

The proposed framework accounts for the uncertainty of $\mathbf{f}$ through $\check{\sigma}_{f_k}^2(\tilde{\mathbf{x}}_i)$ in (10). Therefore, one can adaptively collect a measurement (or a mini-batch of measurements) when a set of available sensing radio pairs is revealed, with the goal of reducing the uncertainty of $\mathbf{f}$. To this end, the conditional entropy [4] is considered as an uncertainty measure of $\mathbf{f}$ at time slot τ , namely,

Algorithm 1 Field estimation via variational Bayes

Input: $\tilde{\mathbf{s}}_t, \mathbf{W}_t, \{\check{\mu}_{f_k}^{(0)}(\tilde{\mathbf{x}}_i), q^{(0)}(z_i = k)\}_{k=0}^1 \forall i$, and N_{Iter} .

- 1: Obtain $\check{\sigma}_{f_k}^2(\tilde{\mathbf{x}}_i) \forall i, k$ with (12)
- 2: **for** $\ell = 1$ to N_{Iter} **do**
- 3: Obtain $\check{\mu}_{f_k}^{(\ell)}(\tilde{\mathbf{x}}_i) \forall i, k$ with (13)
- 4: Obtain $q^{(\ell)}(z_i = k) \forall i, k$ with (11)
- 5: **end for**
- 6: Set $q^*(f_i | z_i) = q^{(N_{\text{Iter}})}(f_i | z_i)$ and $q^*(z_i) = q^{(N_{\text{Iter}})}(z_i) \forall i$
- 7: Estimate $\hat{z}_{\text{MAP}, i} = \arg \max_{z_i \in \{0, 1\}} q^*(z_i) \forall i$
- 8: Estimate $\hat{f}_{i, \text{MMSE}} = \check{\mu}_{f_{\hat{z}_{\text{MAP}, i}}}^*(\tilde{\mathbf{x}}_i) \forall i$
- 9: **return** $\hat{\mathbf{f}}_{\text{MMSE}}, q^*(\mathbf{f} | \mathbf{z})$, and $q^*(\mathbf{z})$

Algorithm 2 Adaptive Bayesian CG cartography

Input: $\tilde{\mathbf{s}}_0, \mathbf{W}_0, \{\check{\mu}_{f_k}^{(0)}(\tilde{\mathbf{x}}_i), q^{(0)}(z_i = k)\}_{i, k}$, N_{Iter} , g_0 , and γ .

- 1: **for** $\tau = 0, 1, \dots$ **do**
- 2: Obtain $\hat{\mathbf{f}}_{\text{MMSE}}, q^*(\mathbf{f} | \mathbf{z})$, and $q^*(\mathbf{z})$
via Alg. 1($\tilde{\mathbf{s}}_\tau, \mathbf{W}_\tau, \{\check{\mu}_{f_k}^{(0)}(\tilde{\mathbf{x}}_i), q^{(0)}(z_i = k)\}_{i, k}, N_{\text{Iter}}$)
- 3: Evaluate $\bar{h}(\mathbf{w}^{(n, n')})$ in (P2') $\forall \{n, n'\} \in \mathcal{M}_{\tau+1}$
- 4: Collect $\check{s}_{\tau+1}$ from (n^*, n'^*) with $\max \bar{h}(\mathbf{w}^{(n, n')})$
- 5: Set $\check{\mathbf{s}}_{\tau+1} = [\check{s}_\tau^\top, \check{s}_{\tau+1}]^\top$ and $\mathbf{W}_{\tau+1} = [\mathbf{W}_\tau, \mathbf{w}_{\tau+1}^{(n^*, n'^*)}]$
- 6: **end for**
- 7: Consider arbitrary locations $\{\mathbf{x}, \mathbf{x}'\} \in \mathcal{A}$
- 8: Estimate $\hat{s}(\mathbf{x}, \mathbf{x}')$ via (2) by using $\hat{\mathbf{f}}_{\text{MMSE}}$
- 9: Estimate $\hat{g}(\mathbf{x}, \mathbf{x}')$ via (1) by using g_0, γ , and $\hat{s}(\mathbf{x}, \mathbf{x}')$

$$H_\tau(\mathbf{f} | \mathbf{z}, \check{\mathbf{s}}_\tau; \boldsymbol{\theta}) = \sum_{\mathbf{z}' \in \mathcal{Z}} \int p(\mathbf{z}', \check{\mathbf{s}}'_\tau; \boldsymbol{\theta}) H(\mathbf{f} | \mathbf{z} = \mathbf{z}', \check{\mathbf{s}}_\tau = \check{\mathbf{s}}'_\tau; \boldsymbol{\theta}) d\check{\mathbf{s}}'_\tau$$

where $H(\mathbf{f} | \mathbf{z} = \mathbf{z}', \check{\mathbf{s}}_\tau = \check{\mathbf{s}}'_\tau; \boldsymbol{\theta}) := -\int p(\mathbf{f} | \mathbf{z} = \mathbf{z}', \check{\mathbf{s}}_\tau = \check{\mathbf{s}}'_\tau; \boldsymbol{\theta}) \ln p(\mathbf{f} | \mathbf{z} = \mathbf{z}', \check{\mathbf{s}}_\tau = \check{\mathbf{s}}'_\tau; \boldsymbol{\theta}) d\mathbf{f}$. To obtain $\check{s}_{\tau+1}$, one can choose a pair of sensors (n^*, n'^*) , or equivalently find $\mathbf{w}^{(n^*, n'^*)}$ minimizing $H_{\tau+1}$. Given $\check{\mathbf{s}}_\tau$, then, $\mathbf{w}^{(n^*, n'^*)}$ can be obtained by solving (see [16] for the derivation of (P2))

$$(P2) \quad \max_{\substack{\mathbf{w}^{(n, n')}: \\ (n, n') \in \mathcal{M}_{\tau+1}}} \mathbb{E}_{p(\mathbf{z} | \check{\mathbf{s}}_\tau; \boldsymbol{\theta})} [h(\mathbf{z}, \mathbf{w}^{(n, n')}; \boldsymbol{\theta})]$$

where $\mathcal{M}_\tau := \{(n, n') | \exists (\mathbf{x}_n - \mathbf{x}_{n'}) \text{ at } \tau, (n, n') \in \{1, \dots, N\}\}$ is a set of available sensing radio pairs at time slot τ and $h(\mathbf{z}, \mathbf{w}; \boldsymbol{\theta}) := \ln(1 + \sigma_{f_k}^{-2} \mathbf{w}^\top \Sigma_{f|z, \check{\mathbf{s}}_\tau; \boldsymbol{\theta}} \mathbf{w})$ with $[\Sigma_{f|z, \check{\mathbf{s}}_\tau; \boldsymbol{\theta}}]_{i, j} := \text{cov}[f_i, f_j | \mathbf{z}, \check{\mathbf{s}}_\tau; \boldsymbol{\theta}] \forall i, j$. However, (P2) cannot be directly solved because $p(\mathbf{z} | \check{\mathbf{s}}_\tau; \boldsymbol{\theta})$ is not available via marginalization of the complex form (6); and evaluating the cost of (P2) is $|\mathcal{Z}| = 2^{N_g}$, which is intractable for large N_g .

Fortunately, the approximate formulation of (P2) can be found by using the variational distributions as follows

$$(P2') \quad \max_{\substack{\mathbf{w}^{(n, n')}: \\ (n, n') \in \mathcal{M}_{\tau+1}}} \underbrace{\sum_{i=1}^{N_g} \mathbb{E}_{q(z_i)} \left[\ln \left(1 + \frac{\check{\sigma}_{f_k}^2(\tilde{\mathbf{x}}_i)}{\sigma_{f_k}^2} w^2(\mathbf{x}_n, \mathbf{x}_{n'}, \tilde{\mathbf{x}}_i) \right) \right]}_{=: \bar{h}(\mathbf{w}^{(n, n')})}$$

Therefore, $\check{s}_{\tau+1}$ can be collected from the pair of sensors (n^*, n'^*) associated with $\mathbf{w}^{(n^*, n'^*)}$ obtained by solving (P2')

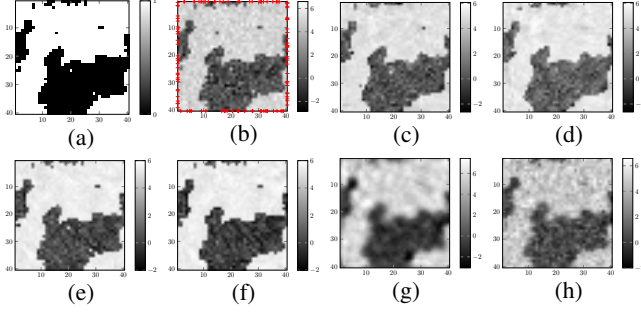


Fig. 1: True fields for synthetic tests: (a) hidden label field $\mathbf{Z}_0$ and (b) spatial loss field $\mathbf{F}_0$ with $N = 120$ sensor locations marked with crosses. Estimated SLFs $\hat{\mathbf{F}}$ at $\tau = 15$ (700 measurements) via: (c) Alg. 2; (d) non-adaptive VB method; (e) adaptive and (f) non-adaptive MCMC methods in [16]; and (g) ridge and (h) TV regularized-LS methods.

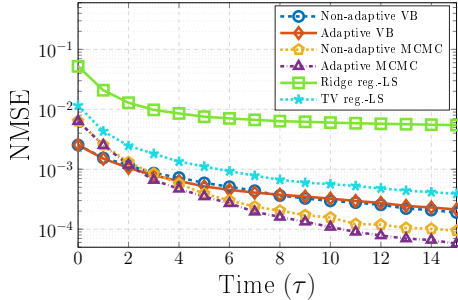


Fig. 2: Progression of channel-gain estimation error.

in a greedy fashion. Note that the proposed data-adaptive sensor selection scheme can be easily extended to a mini-batch setup of size N_{Batch} per time slot τ by finding weight vectors $\{\mathbf{w}^{(n^{(m)}, n'^{(m)})}\}_{m=1}^{N_{\text{Batch}}}$ associated with the N_{Batch} largest values of $\bar{h}(\mathbf{w}^{(n^{(m)}, n'^{(m)})})$ in (P2'), and collecting $\{\tilde{s}_{\tau+1}^{(m)}\}_{m=1}^{N_{\text{Batch}}}$ from pairs of sensors revealed from those weight vectors.

The overall scheme for adaptive Bayesian CG cartography is tabulated as Alg. 2.

4. NUMERICAL TESTS

This section validates the proposed algorithm through synthetic tests. Tomographic measurements were taken by $N = 120$ sensors uniformly deployed on boundaries of $\mathcal{A} := [0.5, 40.5] \times [0.5, 40.5]$, from which the SLF defined over a grid $\{\tilde{\mathbf{x}}_i\}_{i=1}^{1,600} := \{1, \dots, 40\}^2$ was reconstructed. To generate the ground-truth SLF $\mathbf{f}_0$, the hidden label field $\mathbf{z}_0$ was obtained first via the Metropolis algorithm [19] by using the prior of $\mathbf{z}$ in (5) with $\beta = 1.3$. Afterwards, $\mathbf{f}_0$ was constructed to have $f(\tilde{\mathbf{x}}_i) \sim \mathcal{N}(0.2, 1) \forall \tilde{\mathbf{x}}_i \in \mathcal{A}_0$ and $f(\tilde{\mathbf{x}}_j) \sim \mathcal{N}(5, 0.2) \forall \tilde{\mathbf{x}}_j \in \mathcal{A}_1$, resulting in $\boldsymbol{\theta}_f = [0.2, 5, 1, 0.2]^\top$, based on labels in $\mathbf{z}_0$. True $\mathbf{Z}_0 := \text{unvec}(\mathbf{z}_0)$ and $\mathbf{F}_0 := \text{unvec}(\mathbf{f}_0)$ are depicted in Figs. 1a and 1b with sensor locations marked with crosses. Under the mini-batch operation, a set of measurements $\{\tilde{s}_{\tau}^{(m)}\}_{m=1}^{N_{\text{Batch}}}$ per time slot τ were generated via (2) with $\mathbf{f}_0$, w set to the model in (3) with $\lambda = 0.39$, and $\nu_{\tau} \sim \mathcal{N}(0, 0.05)$. To construct $\mathcal{M}_{\tau+1}$ at time

slot τ , $|\mathcal{M}_{\tau+1}| = 100$ pairs of sensors were uniformly selected at random with replacement. Then, $N_{\text{Batch}} = 40$ measurements were collected at step 4 of Alg. 2 for $\tau = 1, \dots, 15$. Note that Alg. 2 replacing steps 3-4 with random sampling was considered as a *non-adaptive VB method* for comparison.

In all synthetic tests, $N_{\text{Iter}} = 10^3$ was used to run the proposed algorithm. While $\{\tilde{\mu}_{f_k}^{(0)}(\tilde{\mathbf{x}}_i)\}_{k=0}^1 \forall i$ were randomly initialized, $\{q^{(0)}(z_i = k)\}_{k=0}^1$ were set to be equally likely $\forall i$. Vector $\tilde{\mathbf{s}}_0$ was collected from randomly selected 100 pairs of sensors. For competing alternatives, ridge and total variation (TV) regularized-LS methods were considered, in which regularization parameters were found by the L-curve [15, Chapter 26] and the generalized cross-validation [8], respectively. The MCMC methods in [16] were also tested as sample-based counterparts of the proposed field estimators.

Since g_0 and γ are known, obtaining $s(\mathbf{x}, \mathbf{x}')$ amounts to finding $g(\mathbf{x}, \mathbf{x}')$; cf. (1). This suggests adopting a performance metric quantifying the mismatch between $s(\mathbf{x}, \mathbf{x}')$ and $\hat{s}(\mathbf{x}, \mathbf{x}')$, using the normalized mean-square error

$$\text{NMSE} := \frac{\mathbb{E} \left[\int_{\mathcal{A}} (s(\mathbf{x}, \mathbf{x}') - \hat{s}(\mathbf{x}, \mathbf{x}'))^2 d\mathbf{x} d\mathbf{x}' \right]}{\mathbb{E} \left[\int_{\mathcal{A}} s^2(\mathbf{x}, \mathbf{x}') d\mathbf{x} d\mathbf{x}' \right]}$$

where the expectation is over the set $\{\mathbf{x}_n\}_{n=1}^N$ of sensor locations and realizations of $\{\nu_{\tau}\}_{\tau}$. Simulations estimated the expectations by averaging over 20 independent Monte Carlo runs with MATLAB. The integrals are approximated by averaging the integrand over 500 pairs of $(\mathbf{x}, \mathbf{x}')$ chosen independently and uniformly at random over the boundary of $\mathcal{A}$.

Figs. 1c–1h depict reconstructed SLFs $\hat{\mathbf{F}} := \text{unvec}(\hat{\mathbf{f}})$ at $\tau = 15$ via Alg. 2 and the competing alternatives. Comparison with $\hat{\mathbf{F}}$ in Fig. 1d obtained by the non-adaptive VB method demonstrates the effectiveness of the proposed algorithm in identifying object patterns in the SLF with adaptively collected measurements. Furthermore, blurry images of $\hat{\mathbf{F}}$ in Figs. 1g and 1h via the regularized LS showcase the benefit of adopting the two-layer Bayesian model for the SLF.

Fig. 2 compares the NMSE of Alg. 2 with those of the competing alternatives using the settings in Figs. 1c–1h. Evidently, the proposed method achieves the NMSE comparable to that of the MCMC method. Note that execution time was 1.5 (sec) per τ for Alg. 2, while that was 184.6 (secs) for the MCMC method in [16]. Although the MCMC approach is a viable solution for CG cartography as shown in Fig. 2, this signifies a benefit of our algorithm in terms of computational complexity reduction, while maintaining high estimation accuracy.

5. CONCLUDING SUMMARY

This paper developed a novel variational Bayes algorithm for adaptive Bayesian channel-gain cartography, which is capable of constructing maps that provide channel-gain between arbitrary locations in a region of interest with low-computational complexity. Efficacy of the proposed algorithm was further improved by the data-adaptive sensor selection strategy.

6. REFERENCES

- [1] P. Agrawal and N. Patwari, "Correlated link shadow fading in multi-hop wireless networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 9, pp. 4024–4036, Aug. 2009.
- [2] E. Axell, G. Leus, and E. G. Larsson, "Overview of spectrum sensing for cognitive radio," in *Proc. Cognitive Inf. Process.*, 2010, pp. 322–327.
- [3] H. Ayasso and A. Mohammad-Djafari, "Joint NDT image restoration and segmentation using Gauss-Markov-Potts models and variational Bayesian computation," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2265–2277, Sep. 2010.
- [4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, NY: USA: Wiley, 1991.
- [5] E. Dall'Anese, S.-J. Kim, and G. B. Giannakis, "Channel gain map tracking via distributed kriging," *IEEE Trans. Veh. Technol.*, vol. 60, no. 3, pp. 1205–1211, 2011.
- [6] Federal Communications Commission, "FCC 11-131," *Unlicensed Operation in the TV Broadcast Bands*, 2011.
- [7] W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. London: Chapman and Hall, 1996.
- [8] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, May 1979.
- [9] B. R. Hamilton, X. Ma, R. J. Baxley, and S. M. Metchik, "Propagation modeling for radio frequency tomography in wireless networks," *IEEE J. Sel. Topics Sig. Proc.*, vol. 8, no. 1, pp. 55–65, Feb. 2014.
- [10] J. Hammersley and P. Clifford, "Markov field on finite graphs and lattices," 1971, unpublished manuscript.
- [11] D. Higdon, "Spatial applications of Markov chain Monte Carlo for Bayesian inference," Ph.D. dissertation, Dept. Stat., Univ. Washington, Seattle, WA, 1994.
- [12] G. Kail, J.-Y. Tournier, F. Hlawatsch, and N. Dobi-geon, "Blind deconvolution of sparse pulse sequences under a minimum distance constraint: a partially collapsed Gibbs sampler method," *IEEE Trans. Sig. Proc.*, vol. 60, no. 6, pp. 2727–2743, 2012.
- [13] S.-J. Kim, E. Dall'Anese, J. A. Bazerque, K. Rajawat, and G. B. Giannakis, "Advances in spectrum sensing and cross-layer design for cognitive radio networks," in *Academic Press Library in Signal Processing: Communications and Radar Signal Processing: 2*. Academic Press, 2013, ch. 9, pp. 471–497.
- [14] S.-J. Kim, E. Dall'Anese, and G. B. Giannakis, "Co-operative spectrum sensing for cognitive radios using kriged Kalman filtering," *IEEE J. Sel. Topics Sig. Process.*, vol. 5, no. 1, pp. 24–36, Feb. 2011.
- [15] C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems*. Philadelphia, PA: SIAM, 1974.
- [16] D. Lee, D. Berberidis, and G. B. Giannakis, "Adaptive Bayesian radio tomography," *IEEE Trans. Sig. Proc.*, Apr. 2019, to appear.
- [17] D. Lee, S.-J. Kim, and G. B. Giannakis, "Channel gain cartography for cognitive radios leveraging low rank and sparsity," *IEEE Trans. Wireless Commun.*, vol. 16, no. 9, pp. 5953–5966, Nov. 2017.
- [18] D. MacKay, "Information-based objective functions for active data selection," *Neural Comput.*, vol. 4, no. 4, pp. 590–604, 1992.
- [19] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chem. Phys.*, vol. 21, no. 6, pp. 1087–1092, Jun. 1953.
- [20] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA: MIT Press, 2012.
- [21] M. Opper and D. Saad, *Advanced mean field methods: Theory and practice*. Cambridge, MA: MIT Press, 2001.
- [22] N. Patwari and P. Agrawal, "Effects of correlated shadowing: Connectivity, localization, and RF tomography," in *Int. Conf. Info. Process. Sensor Networks*, St. Louis, MO, Apr. 2008, pp. 82–93.
- [23] D. Romero, D. Lee, and G. B. Giannakis, "Blind radio tomography," *IEEE Trans. Sig. Proc.*, vol. 66, no. 8, pp. 2055–2069, Apr. 2018.
- [24] D. Smith and M. Smith, "Estimation of binary Markov random fields using Markov chain Monte Carlo," *J. Comput. Graph. Stats.*, vol. 15, no. 1, pp. 207–227, 2006.
- [25] J. Wilson and N. Patwari, "Radio tomographic imaging with wireless networks," *IEEE Trans. Mobile Comput.*, vol. 9, no. 5, pp. 621–632, Jan. 2010.
- [26] Q. Zhao and B. M. Sadler, "A survey of dynamic spectrum access," *IEEE Sig. Process. Mag.*, vol. 24, no. 3, pp. 79–89, 2007.