MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Terahertz Imaging of Binary Reflectance with Variational Bayesian Inference

Fu, H.; Wang, P.; Koike-Akino, T.; Orlik, P.V.; Chi, Y.

TR2018-004 April 2018

Abstract

In this paper, we propose a Bayesian inference approach to extract the binary reflectance pattern of samples from compressed measurements in the terahertz (THz) frequency band. Compared with existing compressed THz imaging methods relying on the sparsity of the reflectance pattern, the proposed Bayesian approach exploits the non-negative binary nature of the reflectance without any assumption on its spatial pattern information and enables a pixel-wise iterative inference approach for fast signal recovery. Numerical evaluation confirms the effectiveness of the proposed approach.

IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)

TERAHERTZ IMAGING OF BINARY REFLECTANCE WITH VARIATIONAL BAYESIAN INFERENCE

Haoyu Fu[2], *Pu Wang*[1], *Toshiaki Koike-Akino*[1], *Philip V. Orlik*[1], *and Yuejie Chi*[2]

[1] Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA 02139, USA
E-mails: {pwang, koike, porlik}@merl.com
[2] Ohio State University, Columbus, OH 43210, USA
E-mails: {fu.436, chi.97}@osu.edu

ABSTRACT

In this paper, we propose a Bayesian inference approach to extract the binary reflectance pattern of samples from compressed measurements in the terahertz (THz) frequency band. Compared with existing compressed THz imaging methods relying on the sparsity of the reflectance pattern, the proposed Bayesian approach exploits the non-negative binary nature of the reflectance without any assumption on its spatial pattern information and enables a pixel-wise iterative inference approach for fast signal recovery. Numerical evaluation confirms the effectiveness of the proposed approach.

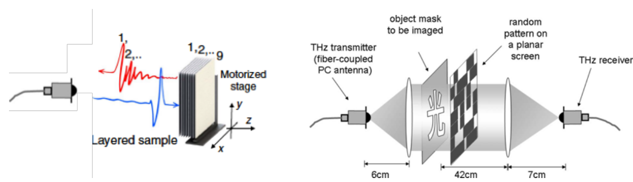
Index Terms— Terahertz sensing, compressed measurements, binary reflectance, variational Bayesian inference.

1. INTRODUCTION

Over the past two decades, there have been increased interests in terahertz (THz) sensing using the time-domain spectroscopy (TDS) in either a reflection or transmission mode, due to the broad applications in gas sensing, moisture analysis, non-destructive evaluation, biomedical diagnosis, package inspection, and security screening [1]. By sending an ultra-short pulse (e.g., 1-2 picoseconds), the THz-TDS system is able to inspect not only the top surface of the sample but also its internal structure, either a defect underneath the top layer or a multi-layer structure, due to its capability of penetrating a wide range of non-conducting materials. At the same time, the ultra-short pulse also gives rise to ultra-wideband spectrum over a band of several THz, providing a spectroscopic inspection of material properties of the sample.

The THz-TDS can operate in a *raster* or *compressed* scanning mode [2–5]. In the raster scanning mode, as shown in Fig. 1 (a), the sample under inspection is illuminated by a THz-TDS point source with a time-compact source pulse and a small spot size (or aperture). The THz-TDS emitter sends a focused beam at a normal incident angle to inspect a small area (or a pixel) of the sample, the detector then samples corresponding reflected waveform via the electro-optic sampling process, and a programmable mechanical raster moves the sample in the plane perpendicular to the incidental waveform in order to measure the two-dimensional surface of the sample. The THz-TDS with the raster scanning mode has already been commercialized with a fast scanning rate (up to 1,000 Hz) and applied to, among other industrial applications [2], art and archaeology [3], quality control [6], thickness estimation [7–9] and multi-layer content extraction [10–12]. One of key challenges is to address the depth variations and its induced delay/phase variation from one pixel to another due to either the irregular sample surface or the vibration from the mechanical scanning process.

The work of H. Fu was performed during his internship at MERL.

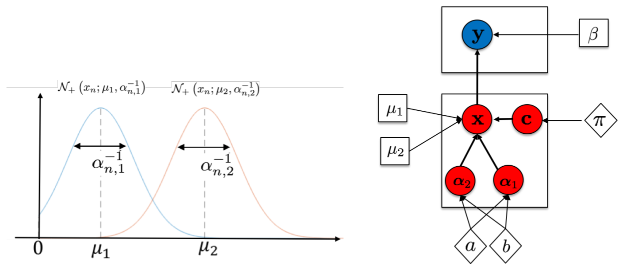


(a) (b)

Fig. 1. THz-TDS sensing with a) a raster scanning (from [11]) and b) a compressed scanning (from [13]).

In the compressed scanning mode, as shown in Fig. 1 (b), the THz pulse is first collimated to a broad beam and then spatially encoded with a random mask with the help of a spatial light modulator (SLM) that operates in the terahertz regime [13, 14]. At the receiver side, the spatially encoded beam is re-focused by a focusing lens and received by a single-pixel photoconductive detector [13–15]. In other words, only one measurement is formed for a mask at a time. The compressed scanning process repeats with different realizations of random masks and collects multiple sequential measurements. The sample image can then be recovered by, normally, sparsity-driven minimization methods. In [13], the total-variation minimization method was used to reconstruct the sample image of a Chinese character, "light", with a small number of measurements than the number of pixels, as shown in Fig. 1 (b).

In this paper, rather than relying on the sparsity assumption of the sample spatial pattern, we here exploit only the *non-negative binary* nature of reflectance coefficient of the sample and recover its reflectance pattern with compressed measurements. This is motivated by applications such as absolute positioning encoder systems where a non-sparse binary pseudo-random pattern (e.g., quick response (QR) code) may be used for the sample. To this end, the proposed method imposes a hierarchical truncated Gaussian mixture prior model to enforce the non-negative binary feature of the reflectance, and uses the principles of generalized approximate message passing (GAMP) and variational Bayesian inference to develop a decoupled pixel-wise iterative recovery algorithm for fast signal recovery. The key challenge here is that, to update the deterministic unknown parameters, i.e., the two unknown means of reflectance coefficients, we need to compute the expectation of the logarithm of two normalization factors (due to the truncated Gaussian mixture model) over the posterior distribution, resulting in no closed-form expressions. To address this issue, we propose an approximate, closed-form updating rule by replacing the expectations with



(a) (b)

Fig. 2. (a) The truncated Gaussian mixture prior distribution of the nth pixel coefficient x_n with pixel-dependent variances $\{\alpha_{n,1}^{-1}, \alpha_{n,2}^{-1}\}_{n=1}^N$ and two shared means $\{\mu_1, \mu_2\}$; (b) A graphical representation of the proposed hierarchical prior model.

its values from the previous iteration. The performance is numerically evaluated by using the Monte-Carlo simulation on a sample with a binary QR-like reflectance pattern.

2. SIGNAL MODEL

Let $\mathbf{x} = \left[x_1, x_2, \ldots, x_N\right]^{\mathrm{T}}$ denote a binary reflectance vector by stacking the columns of the two-dimensional reflectance matrix of the sample. As the THz source illuminates the sample from a spatially encoded mask, the received measurement can be expressed as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}, \quad x_n \in \{\mu_1, \mu_2\}, \tag{1}$$

where $\mathbf{A} = \left[\mathbf{a}_1, \ldots, \mathbf{a}_M\right]^{\mathrm{T}}$ is the measurement matrix, $\mathbf{v} = \left[v_1, \ldots, v_M\right]^{\mathrm{T}}$ is the Gaussian distributed noise with zero mean and an unknown variance β^{-1}, i.e., $\mathbf{v} \sim \mathcal{N}(\mathbf{0}, \beta^{-1}\mathbf{I}_M)$, $\mathbf{y} = \left[y_1, \ldots, y_M\right]^{\mathrm{T}}$, M is the number of measurements, and μ_i for $i = 1, 2$ are two unknown reflectance coefficients. Moreover, the reflectance coefficient $\mathbf{x}$ is assumed to be non-negative, i.e., $x_n \geq 0$. The signal model of (1) can, in fact, describe both raster and compressed scanning acquisitions:

- In the case of the *raster scanning*, i.e., each pixel is illuminated and measured individually, we have $M = N$ and $\mathbf{A}$ reduces to a diagonal matrix with diagonal elements responsible for the depth variation [12, Section III.I.4].
- In the case of the *compressed scanning*, e.g., the single-pixel THz camera [13], we have $M < N$ and each row of the measurement matrix $\mathbf{A}$ corresponds to one random mask pattern used to form one measurement y_m.

To account for the non-negative binary feature of $\mathbf{x}$, we introduce the following hierarchical Gaussian mixture prior distribution,

$$\begin{aligned} & p\left(x_n | \alpha_{n,1}, \alpha_{n,2}, c_n; \mu_1, \mu_2\right) \\ & = \mathcal{N}_+\left(x_n; \mu_1, \alpha_{n,1}^{-1}\right)^{c_n} \cdot \mathcal{N}_+\left(x_n; \mu_2, \alpha_{n,2}^{-1}\right)^{1-c_n}, \end{aligned} \tag{2}$$

where $c_n \in \{0, 1\}$ is a binary latent label variable for the pixel x_n, and the truncated Gaussian distribution is given as

$$\mathcal{N}_+\left(x; \mu, \alpha^{-1}\right) = \begin{cases} \eta^{-1}\sqrt{\frac{\alpha}{2\pi}} \exp\left(-\frac{\alpha(x-\mu)^2}{2}\right), & x \geq 0, \\ 0, & x < 0, \end{cases} \tag{3}$$

with μ as its mean, α^{-1} as the variance (or α as the precision parameter) and $\eta = 1 - \Phi\left(-\mu\sqrt{\alpha}\right)$ as the normalization factor where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. In addition to (2), the binary label vector $\mathbf{c} = [c_1, \ldots, c_N]^{\mathrm{T}}$ follows an *i.i.d.* Bernoulli distribution with parameter π,

$$p\left(c_n; \pi\right) = (\pi)^{c_n}\left(1 - \pi\right)^{1-c_n}. \tag{4}$$

With both (2) and (4), we can show that the pixel-wise reflectance coefficient x_n has independent truncated Gaussian mixture prior distribution by integrating over the latent label variable c_n

$$\begin{aligned} & p\left(x_n | \alpha_{n,1}, \alpha_{n,2}; \mu_1, \mu_2\right) \\ & = \sum_{c_n \in \{0,1\}} p\left(x_n | \alpha_{n,1}, \alpha_{n,2}, c_n; \mu_1, \mu_2\right) p\left(c_n; \pi\right) \\ & = \pi\mathcal{N}_+\left(x_n; \mu_1, \alpha_{n,1}^{-1}\right) + (1 - \pi)\mathcal{N}_+\left(x_n; \mu_2, \alpha_{n,2}^{-1}\right). \end{aligned} \tag{5}$$

The resulting truncated Gaussian mixture prior distribution of x_n is illustrated in Fig. 2 (a) with *pixel-dependent* precision parameters, i.e., $\alpha_{n,1}$ and $\alpha_{n,2}$, and two *shared* mean parameters μ_1 and μ_2.

Furthermore, we treat the pixel-dependent precision parameters $\boldsymbol{\alpha}_1 = [\alpha_{1,1}, \ldots, \alpha_{N,1}]^{\mathrm{T}}$ and $\boldsymbol{\alpha}_2 = [\alpha_{1,2}, \ldots, \alpha_{N,2}]^{\mathrm{T}}$ as *i.i.d.* random variables and assume the Gamma distribution as their hyperprior distribution

$$p\left(\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2; a, b\right) = \prod_{i=1}^{2}\prod_{n=1}^{N} \mathrm{Gamma}\left(\alpha_{n,i} | a, b\right), \tag{6}$$

where $\mathrm{Gamma}\left(\alpha | a, b\right) = \Gamma\left(a\right)^{-1} b^a \alpha^{a-1} e^{-b\alpha}$ with $a = b = 10^{-6}$ for non-informative hyperpriors on $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$. Overall, the hierarchical truncated Gaussian mixture model can be described in a graphical representation shown in Fig. 2 (b), where blue and red circles denote observed and hidden random variables, respectively, squares denote the unknown deterministic model parameters, and diamonds denote the pre-determined user parameters (π, a and b).

3. PROPOSED APPROACH

In this section, we derive a specialized variational Bayesian inference for the posterior distribution of the hidden random variables and a cost function to update the deterministic model parameters. Particularly, a two-step approach is used: First, we factorize the original likelihood function, coupled over $\mathbf{x}$ due to the measurement matrix $\mathbf{A}$, into a pixel-wise decoupled likelihood function with the principle of GAMP. Second, with the decoupled likelihood function on $\mathbf{x}$, the variational expectation-maximization (EM) algorithm is used to derive the posterior distribution and the Q-function to update the unknown model parameters.

3.1. Pixel-Wise Decoupled Likelihood Function

The likelihood function of $\mathbf{y}$ is given by

$$p\left(\mathbf{y} | \mathbf{x}; \beta\right) = \left(2\pi\beta^{-1}\right)^{-M/2} e^{-\frac{\beta\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2^2}{2}}, \tag{7}$$

where each measurement y_m is coupled with all pixels $\{x_n\}_{n=1}^N$. In order to enable a fast, pixel-wise Bayesian inference, we can approximate the likelihood function of (7) onto the pixel coefficient x_n:

$$p\left(\mathbf{y} | \mathbf{x}; \beta\right) \approx \prod_{n=1}^{N} p(x_n | \hat{r}_n, \hat{\tau}_n) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\hat{\tau}_n}} e^{-\frac{(x_n - \hat{r}_n)^2}{2\hat{\tau}_n}}. \tag{8}$$

In other words, the approximated marginal likelihood function is given by $x_n \sim \mathcal{N}\left(\hat{r}_n, \hat{\tau}_n\right)$ where the approximated mean $\hat{r}_n$ and variance $\hat{\tau}_n$ can be found by the GAMP algorithm. As a result, the likelihood function of $\mathbf{y}$ is factorized as a product of independent decoupled likelihood function of x_n with mean $\hat{r}_n$ and variance $\hat{\tau}_n$ (refer to the Appendix). It is worth noting that an analogous decoupling process of (8) has been used in wireless multi-antenna communications [16–19].

3.2. Variational Bayesian Inference

Given the decoupled likelihood function of (8), we use the variational Bayesian framework [20] to derive the posterior distributions of all hidden random variables $\mathbf{z} = \{\mathbf{x}, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathbf{c}\}$ (red circles in Fig. 2), and then update the unknown deterministic parameters $\boldsymbol{\theta} = \{\beta, \mu_1, \mu_2\}$ (squares in Fig. 2) by maximizing the expectation of the complete likelihood function over the posterior distribution of the hidden variables.

3.2.1. Posterior distributions of hidden variables $\{\mathbf{x}, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathbf{c}\}$

In the conventional Bayesian framework, the posterior of the hidden variables can be found via the E-step of the EM framework. Generally, the E-step is to find a probability density function $q(\mathbf{z})$ which, given the current estimate of the model parameters θ, maximizes the marginal likelihood of the measurement $p(\mathbf{y}; \boldsymbol{\theta})$. With the variational Bayesian framework, we can factorize $q(\mathbf{z}) \approx q(\mathbf{x})q(\boldsymbol{\alpha}_1)q(\boldsymbol{\alpha}_2)q(\mathbf{c})$ and, instead of joint optimization over $\mathbf{z}$, the E-step can find the optimal probability density function of each class of hidden variables, leading to

$$\ln q(\mathbf{x}) = \langle \ln p(\mathbf{y}, \mathbf{z}; \boldsymbol{\theta}) \rangle_{q(\boldsymbol{\alpha}_1)q(\boldsymbol{\alpha}_2)q(\mathbf{c})} + \text{const}, \tag{9}$$

$$\ln q(\boldsymbol{\alpha}_1) = \langle \ln p(\mathbf{y}, \mathbf{z}; \boldsymbol{\theta}) \rangle_{q(\mathbf{x})q(\boldsymbol{\alpha}_2)q(\mathbf{c})} + \text{const}, \tag{10}$$

$$\ln q(\boldsymbol{\alpha}_2) = \langle \ln p(\mathbf{y}, \mathbf{z}; \boldsymbol{\theta}) \rangle_{q(\mathbf{x})q(\boldsymbol{\alpha}_1)q(\mathbf{c})} + \text{const}, \tag{11}$$

$$\ln q(\mathbf{c}) = \langle \ln p(\mathbf{y}, \mathbf{z}; \boldsymbol{\theta}) \rangle_{q(\mathbf{x})q(\boldsymbol{\alpha}_1)q(\boldsymbol{\alpha}_2)} + \text{const}, \tag{12}$$

where $p(\mathbf{y}, \mathbf{z}) = p(\mathbf{y}, \mathbf{x}, \boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \mathbf{c}; \boldsymbol{\theta})$ is the complete likelihood function of the observable and hidden variables and $q(\cdot)$ is the posterior distribution of the corresponding class of hidden variables.

We start with the first class of hidden variables: the pixel-wise reflectance coefficient $\mathbf{x}$. By keeping terms related to x_n in (2) and (8) in (9), we can show that $\{x_n\}_{n=1}^N$ have independent truncated Gaussian posterior distributions

$$q(x_n) = \begin{cases} \phi_n^{-1} \frac{1}{\sqrt{2\pi}\tilde{\sigma}_n} \exp\left(-\frac{(x_n - \tilde{\mu}_n)^2}{2\tilde{\sigma}_n^2}\right), & x_n \geq 0, \\ 0, & x_n < 0, \end{cases} \tag{13}$$

where the posterior mean $\tilde{\mu}_n$ and posterior variance $\tilde{\sigma}_n^2$ are given as

$$\tilde{\sigma}_n^2 = \left(\langle c_n \rangle \langle \alpha_{n,1} \rangle + \langle 1 - c_n \rangle \langle \alpha_{n,2} \rangle + 1/\hat{\tau}_n\right)^{-1}, \tag{14}$$

$$\tilde{\mu}_n = \left(\langle c_n \rangle \langle \alpha_{n,1} \rangle \mu_1 + \langle 1 - c_n \rangle \langle \alpha_{n,2} \rangle \mu_2 + \hat{r}_n/\hat{\tau}_n\right) \tilde{\sigma}_n^2, \tag{15}$$

with $\phi_n = 1 - \Phi(-\tilde{\mu}_n/\tilde{\sigma}_n)$ as the normalization factor.

For the second class of hidden variables of $\boldsymbol{\alpha}_1$, its posterior distribution is the Gamma distribution with the help of (2), (6) and (10)

$$q(\alpha_{n,1}) = \text{Gamma}(\alpha_{n,1} | \tilde{a}_{n,1}, \tilde{b}_{n,1}), \tag{16}$$

with $\tilde{a}_{n,1} = a + 0.5\langle c_n \rangle$ and $\tilde{b}_{n,1} = b + 0.5\langle c_n \rangle \langle (x_n - \mu_1)^2 \rangle$.

Similarly, for the third class of $\boldsymbol{\alpha}_2$, its posterior distribution is also the Gamma distribution

$$q(\alpha_{n,2}) = \text{Gamma}(\alpha_{n,2} | \tilde{a}_{n,2}, \tilde{b}_{n,2}), \tag{17}$$

Algorithm 1 THz Imaging Algorithm of Binary Reflectance

Initialization: set the iteration counter as $k = 1$

- Set $\pi = 0.5$ and $a = b = 10^{-5}$;
- Set posterior quantities of hidden variables as $\langle x_n \rangle = 0$, $\langle x_n^2 \rangle = 1$, $\langle \alpha_{n,i} \rangle_{i=1}^2 = 1$, $\langle c_n \rangle = 0.5$;
- Set model parameters as $\beta = 10^3$ and $\mu_1 = \mu_2 = 0.5$;

Repeat: until a stopping rule is met

1. Compute the decoupled likelihood function of (8) according to the Appendix;
2. Update $q(x_n)$, $\langle x_i \rangle$ and $\langle x_i^2 \rangle$ according to (14) and (15);
3. Update $q(\alpha_{n,i})$, $\langle \alpha_{n,i} \rangle$, $\langle \ln \alpha_{n,i} \rangle$, $i = \{1, 2\}$, according to (16) and (17);
4. Update $q(c_n)$ and $\langle c_n \rangle$ according to (18);
5. Update β according to (20);
6. Update μ_1 and μ_2 according to (22) and (23);
7. Increase $k = k + 1$ and return to Step 1;

---

with $\tilde{a}_{n,2} = a + 0.5\langle 1 - c_n \rangle$ and $\tilde{b}_{n,2} = b + 0.5\langle 1 - c_n \rangle \langle (x_n - \mu_2)^2 \rangle$.

Finally, for the latent label variable $\mathbf{c}$, its posterior distribution is the Bernoulli distribution with the help of (2), (4) and (12)

$$\ln q(c_n) = (l_{n,1} - l_{n,2})c_n + \text{const}, \tag{18}$$

with $l_{n,1} = 0.5\langle \ln \alpha_{n,1} \rangle - 0.5\langle \alpha_{n,1} \rangle \langle (x_n - \mu_1)^2 \rangle - \langle \ln \eta_{n,1} \rangle + \ln \pi$, $l_{n,2} = 0.5\langle \ln \alpha_{n,2} \rangle - 0.5\langle \alpha_{n,2} \rangle \langle (x_n - \mu_2)^2 \rangle - \langle \ln \eta_{n,2} \rangle + \ln(1 - \pi)$.

To compute the above parameters associated with the posterior distributions, we need the following expressions:

$$\langle x_n \rangle = \tilde{\mu}_n + \tilde{\sigma}_n \cdot \phi(-\tilde{\mu}_n/\tilde{\sigma}_n)/\phi_n, \quad \langle x_n^2 \rangle = \tilde{\sigma}_n^2 + \tilde{\mu}_n \cdot \langle x_n \rangle,$$
$$\langle \alpha_{n,i} \rangle = \tilde{a}_{n,i}/\tilde{b}_{n,i}, \quad \langle \ln \alpha_{n,i} \rangle = \psi(\tilde{a}_{n,i}) - \ln \tilde{b}_{n,i}, \quad i = 1, 2,$$
$$\langle c_n \rangle = (1 + e^{l_{n,2} - l_{n,1}})^{-1},$$

where $\psi(a) = \frac{\partial}{\partial a} \ln \Gamma(a)$ is the digamma function [21].

3.2.2. Updating for deterministic parameters $\{\beta, \mu_1, \mu_2\}$

The next step is to find an updating rule for the deterministic unknown parameters by maximizing the following Q-function [20]

$$\left\{\boldsymbol{\theta}^{(k+1)}\right\} = \max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(k)}) = \underset{\boldsymbol{\theta}}{\text{argmin}} \langle \ln p(\mathbf{y}, \mathbf{z}; \boldsymbol{\theta}) \rangle_{q(\mathbf{z})}. \tag{19}$$

First we use (19) to derive the updating rule for the noise variance β^{-1}, which reduces to

$$\left(\beta^{-1}\right)^{(k+1)} = \frac{\sum_{m=1}^M \langle (y_m - w_m)^2 \rangle}{M}, \tag{20}$$

where w_m is the m-th element of $\mathbf{w} = \mathbf{A}\mathbf{x}$ whose posterior can be found in Appendix.

Then we obtain the updating rule for the two shared means μ_1 and μ_2. With the above derivations, the corresponding Q-function reduces to the function $g(\mu_1, \mu_2)$ defined as

$$g(\mu_1, \mu_2) = \sum_{n=1}^N \Big[\langle c_n \rangle (\langle \ln \eta_{n,1} \rangle - 0.5\langle \alpha_{n,1} \rangle (\mu_1^2 - 2\langle x_n \rangle \mu_1))$$
$$- \langle 1 - c_n \rangle (\langle \ln \eta_{n,2} \rangle - 0.5\langle \alpha_{n,2} \rangle (\mu_2^2 - 2\langle x_n \rangle \mu_2)) \Big], \tag{21}$$

where the two normalization factors $\eta_{n,i} = 1 - \Phi\left(-\mu_i \sqrt{\alpha_{n,i}}\right)$, $i = \{1, 2\}$ are a function of the hidden variables $\{\mu_i\}_{i=1}^2$ and $\{\alpha_{n_i}\}_{i=1}^2$. As a result, we need to compute the expectation of $\ln \eta_{n,1}$ and $\ln \eta_{n,2}$ over the posterior distributions of these hidden variables which results in no closed-form expressions. Instead, we replace

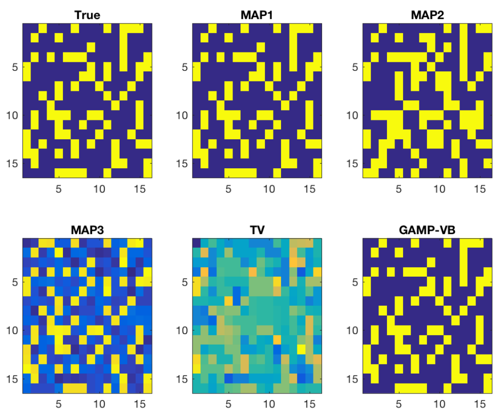


Fig. 3. Recovered images of a QR-like pseudo random pattern with binary reflectance at 0.3 and 0.8.

$\langle \ln \eta_{n,1} \rangle$ and $\langle \ln \eta_{n,2} \rangle$ in (21) by their current estimates from the previous iteration, i.e., $\ln \eta_{n,1}^{(k)}$ and $\ln \eta_{n,2}^{(k)}$. With this approximation, the updates of μ_1 and μ_2 are decoupled as

$$\mu_1^{(k+1)} = \frac{\sum_{n=1}^{N} \langle c_n \rangle \langle \alpha_{n,1} \rangle \langle x_n \rangle}{\sum_{n=1}^{N} \langle c_n \rangle \langle \alpha_{n,1} \rangle}, \tag{22}$$

$$\mu_2^{(k+1)} = \frac{\sum_{n=1}^{N} \langle 1 - c_n \rangle \langle \alpha_{n,2} \rangle \langle x_n \rangle}{\sum_{n=1}^{N} \langle 1 - c_n \rangle \langle \alpha_{n,2} \rangle}, \tag{23}$$

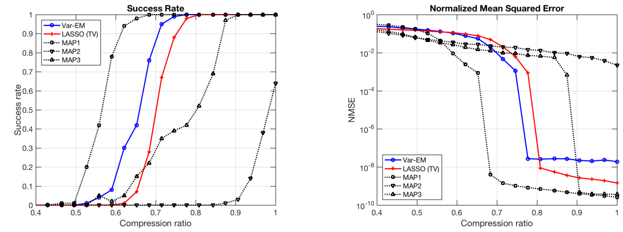
which turn out to be the weighted averages of all posterior means.

Overall, the implementation of the proposed method is described in Algorithm 1 where a stopping rule can be either the number of iterations and/or the difference between two consecutive iterations.

4. NUMERICAL RESULTS

In this section, numerical results are provided to compare the proposed THz imaging method with other existing approaches in terms of the success recovery rate and the normalized mean squared error (NMSE) as a function of a compression ratio M/N. Specifically, we consider 1) the maximum *a posteriori* (MAP) approach [12] with an extension of the decoupled likelihood function in Section 3.1 to the underdetermined scenario of $M/N < 1$, and 2) the total variation minimization approach of [13]. Since the MAP approach requires a preset prior means and variances, we consider 3 implementations: 1) "MAP1" with true means and small variances; 2) "MAP2" with wrong means and small variances; and 3) "MAP3" with true means but large variances. As shown in Fig. 3, with $M/N = 0.7$, the MAP1 provides a recovered image, almost identical to the ground truth. On contrary, the MAP2 with wrong means produces a binary image which is deviated to the ground truth and the MAP3 gives a non-binary image due to the large variances used in the algorithm. Meanwhile, the total variation minimization appears to fail in this case as the spatial isotropic gradients are not sparse. Finally, the proposed variational Bayesian approach (denoted as "GAMP-VB") gives a similar image to the ground truth as well as the MAP1.

We then evaluate NMSE performance for randomly generated QR-like patterns in the Monte-Carlo simulation. In each Monte-Carlo trial, the measurement matrix is generated as the random Gaussian matrix with zero mean and unit variance. And the signal-to-noise ratio (SNR) is defined as $\beta \|\mathbf{A}\mathbf{x}\|_2^2/M$. The NMSE metric is defined as $\|\hat{\mathbf{x}} - \mathbf{x}\|_2^2/\|\mathbf{x}\|_2^2$. The Monte-Carlo trial is considered



(a) (b)

Fig. 4. Performance comparison in terms of (a) the success rate and (b) NMSE as a function of the compression ratio M/N.

to be a success if NMSE $\leq 10^{-3}$. Fig. 4 (a) shows that the success rate as a function of the compression ratio. It is seen that the proposed variational Bayesian approach outperforms the total-variation minimization approach and the two MAP implementations (MAP2 and MAP3). The MAP1 with true means and small variances again serves as an upper bound on these considered methods. The measured NMSE versus the compression ratio is shown in Fig. 4 (b) from which similar observations can be made.

5. CONCLUSION

In this paper, a new THz imaging algorithm is proposed which captures the non-negative binary reflection pattern by introducing the hierarchical prior signal model. The signal recovery algorithm has been derived by using the GAMP framework to decouple the likelihood function into the pixel level and using the variational Bayesian framework to update the hidden random variables and unknown deterministic model parameters. It is shown that the proposed algorithm outperforms the total variation minimization approach and the MAP approach for a QR-like binary reflectance pattern.

6. APPENDIX

The approximate likelihood function of (8) can be obtained by using the GAMP framework [22] with inputs from the means $\hat{x}_n = \langle x_n \rangle_{q(x_n)}$, variances $\tau_n^x = \langle (x_n - \hat{x}_n)^2 \rangle_{q(x_n)}$, and the noise variance β^{-1}. Particularly, to compute the decoupled likelihoods $\mathcal{N}(x_n | \hat{r}_n, \hat{\tau}_n)$ and the posterior likelihood of the noiseless measurement $\mathcal{N}(w_m | \hat{w}_m, \hat{\tau}_m^w)$, we follow the steps below:

- Initialize $\hat{s}_m = 0$ for $m = 1, \cdots, M$;
- **Step 1**: for all $m = 1, \cdots, M$:

$$\hat{\tau}_m^x = \sum_n A_{mn}^2 \tau_n^x, \quad \hat{p}_m^x = \sum_n A_{mn} \hat{x}_n - \hat{\tau}_m^x \hat{s}_m,$$

  where A_{mn} is the (m, n)th element of $\mathbf{A}$.
- **Step 2**: for all $m = 1, \cdots, M$, compute the posterior mean and variance of w_m with respect to $p(w_m | y_m, \hat{\tau}_m^x, \hat{p}_m^x)$, i.e.,

$$\hat{w}_m = \langle w_m \rangle_{p(w_m | y_m, \hat{\tau}_m^x, \hat{p}_m^x)},$$
$$\hat{\tau}_m^w = \langle (w_m - \hat{w}_m)^2 \rangle_{p(w_m | y_m, \hat{\tau}_m^x, \hat{p}_m^x)},$$

  and update $\hat{s}_m = (\hat{w}_m - \hat{p}_m^x)/\hat{\tau}_m^x$ and $\hat{\tau}_m^s = (1 - \hat{\tau}_m^w/\hat{\tau}_m^x)/\hat{\tau}_m^x$.
- **Step 3**: for all $n = 1, \cdots, N$, compute the mean and variance of the decoupled likelihood function

$$\hat{\tau}_n = \left(\sum_m A_{mn}^2 \hat{\tau}_m^s \right)^{-1}, \quad \hat{r}_n = \hat{x}_n + \hat{\tau}_n \sum_m A_{mn} \hat{s}_m.$$

7. REFERENCES

[1] N. Horiuchi, "Terahertz technology: Endless applications," *Nature Photonics*, vol. 4, no. 4, pp. 140, Sept. 2010.

[2] D. L. Woolard, R. Brown, M. Pepper, and M. Kemp, "Terahertz frequency sensing and imaging: A time of reckoning future applications?," *Proceedings of the IEEE*, vol. 93, no. 10, pp. 1722–1743, Oct. 2005.

[3] J.B. Jackson, J. Bowen, G. Walker, J. Labaune, G. Mourou, M. Menu, and K. Fukunaga, "A survey of Terahertz applications in cultural heritage conservation science," *IEEE Trans. on Terahertz Science And Technology*, vol. 1, no. 1, pp. 220–231, Sept. 2011.

[4] E. P. Parrott, S. M. Sy, T. Blu, V. P. Wallace, and E. Pickwell-MacPherson, "Terahertz pulsed imaging in vivo: measurements and processing methods," *Journal of Biomedical Optics*, vol. 16, no. 10, pp. 106010–1–106010–8, 2011.

[5] G. C. Walker, J. W. Bowen, J. Labaune, J-B. Jackson, S. Hadjiloucas, J. Roberts, G. Mourou, and M. Menu, "Terahertz deconvolution," *Optics Express*, vol. 20, no. 25, pp. 27230–27241, Dec. 2012.

[6] A. I. Hernandez-Serrano, S. C. Corzo-Garcia, E. Garcia-Sanchez, M. Alfaro, and E. Castro-Camus, "Quality control of leather by terahertz time-domain spectroscopy," *Applied Optics*, vol. 53, no. 33, pp. 7872–7876, Nov. 2014.

[7] L. Duvillaret, F. Garet, and J. L. Coutaz, "A reliable method for extraction of material parameters in terahertz time-domain spectroscopy," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 2, no. 3, pp. 739–746, Sept. 1996.

[8] T. D. Dorney, R. G. Baraniuk, and D. M. Mittleman, "Material parameter estimation with terahertz time-domain spectroscopy," *J. Opt. Soc. Am. A*, vol. 18, no. 7, pp. 1562–1571, July 2001.

[9] R. Wilk, I. Pupeza, R. Cernat, and M. Koch, "Highly accurate THz time-domain spectroscopy of multilayer structures," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 14, no. 2, pp. 392–398, Mar. 2008.

[10] N. Sunaguchi, Y. Sasaki, N. Maikusa, M. Kawai, T. Yuasa, and C. Otani, "Depth-resolving THz imaging with tomosynthesis," *Optics Express*, vol. 17, no. 12, pp. 9558–9570, June 2009.

[11] A. Redo-Sanchez, B. Heshmat, A. Aghasi, S. Naqvi, M. Zhang, J. Romberg, and R. Raskar, "Terahertz time-gated spectral imaging for content extraction through layered structures," *Nature Communications*, vol. 7, pp. 1–7, Sept. 2016.

[12] A. Aghasi, B. Heshmat, A. Redo-Sanchez, J. Romberg, and R. Raskar, "Sweep distortion removal from terahertz images via blind demodulation," *Optica*, vol. 3, no. 7, pp. 754–762, July 2016.

[13] W.L. Chan, K. Charan, D. Takhar, K. F. Kelly, R. G. Baraniuk, and D. M. Mittleman, "A single-pixel terahertz imaging system based on compressed sensing," *Applied Physics Letters*, vol. 93, no. 121105, Sept. 2008.

[14] C. M. Watts, D. Shrekenhamer, J. Montoya, G. Lipworth, J. Hunt, T. Sleasman, S. Krishna, D. R. Smith, and W. J. Padilla, "Terahertz compressive imaging with metamaterial spatial light modulators," *Nature Photonics*, vol. 8, no. 8, pp. 605–609, Aug. 2014.

[15] S. A. N. Saqueb and K. Sertel, "Phase-sensitive single-pixel THz imaging using intensity-only measurements," *IEEE Transactions on Terahertz Science and Technology*, vol. 6, no. 6, pp. 810–816, Nov. 2016.

[16] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang, and Q. Guo, "Low-complexity iterative detection for large-scale multiuser MIMO-OFDM systems using approximate message passing," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 902–915, Oct. 2014.

[17] C. Jeon, R. Ghods, A. Maleki, and C. Studer, "Optimality of large MIMO detection via approximate message passing," in *2015 IEEE International Symposium on Information Theory (ISIT)*, June 2015, pp. 1227–1231.

[18] R. Ghods, C. Jeon, A. Maleki, and C. Studer, "Optimal large-MIMO data detection with transmit impairments," in *The 53rd Annual Allerton Conference on Communication, Control, and Computing*, Oct. 2015.

[19] H. Bao, J. Fang, Z. Chen, H. Li, and S. Li, "An efficient Bayesian PAPR reduction method for OFDM-based massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 4183–4195, 2016.

[20] D. G. Tzikas, A. C. Likas, and N. P. Galatsanos, "The variational approximation for Bayesian inference," *IEEE Signal Processing Magazine*, vol. 25, no. 6, pp. 131–146, 2008.

[21] M. Abramowitz and I. A.Ryzhik, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, Dover, New York, NY, 1972.

[22] S. Rangan, "Generalized approximate message passing for estimation with random linear mixing," in *2011 IEEE International Symposium on Information Theory Proceedings*, July 2011, pp. 2168–2172.