ORIGINAL ARTICLE   **Open Access**

Extension of emission expectation maximization lookalike algorithms to Bayesian algorithms

Gengsheng L. Zeng[1,2*] and Ya Li[3]

Abstract

We recently developed a family of image reconstruction algorithms that look like the emission maximum-likelihood expectation-maximization (ML-EM) algorithm. In this study, we extend these algorithms to Bayesian algorithms. The family of emission-EM-lookalike algorithms utilizes a multiplicative update scheme. The extension of these algorithms to Bayesian algorithms is achieved by introducing a new simple factor, which contains the Bayesian information. One of the extended algorithms can be applied to emission tomography and another to transmission tomography. Computer simulations are performed and compared with the corresponding un-extended algorithms. The total-variation norm is employed as the Bayesian constraint in the computer simulations. The newly developed algorithms demonstrate a stable performance. A simple Bayesian algorithm can be derived for any noise variance function. The proposed algorithms have properties such as multiplicative updating, non-negativity, faster convergence rates for bright objects, and ease of implementation. Our algorithms are inspired by Green's one-step-late algorithm. If written in additive-update form, Green's algorithm has a step size determined by the future image value, which is an undesirable feature that our algorithms do not have.

Keywords: Image reconstruction, Tomography, Iterative reconstruction algorithm

Introduction

This work is inspired by Green's one-step-late (OSL) expectation-maximization (EM) algorithm [1, 2]. Green's algorithm became popular because it is user-friendly and easy to implement. It has a wide range of applications, such as in positron emission tomography (PET) and single photon emission computed tomography (SPECT) [3–7]. Green's algorithm also has applications in other fields, such as the minimization of the penalized *I*-divergence [8]. Furthermore, Green's algorithm may diverge [9]. This study improves Green's algorithm, making it more stable and more applicable for various noise models.

Green's algorithm is a maximum a posterior (MAP) algorithm, using image-domain constraints for emission tomography. Other MAP algorithms exist [10–15]. In

ref. [12], a method of projection onto convex sets (POCS) was proposed to enforce data fidelity, total-variation (TV) minimization, and image non-negativity. In addition, a GPU algorithm was proposed in ref. [13] to combat the long computation time in combined EM and TV minimization. Filtered backprojection (FBP) reconstruction was proposed for use as the initial image for penalized weighted least-squares (PWLS-TV) reconstruction [12]. Furthermore, in ref. [13] monotonic algorithms for transmission tomography penalized likelihood image reconstruction were developed based on paraboloidal surrogate functions. A similar idea using surrogate functions was reported in refs.[16, 17].

Most recently, we developed a family of emission-EM-lookalike algorithms [10]. These were iterative algorithms in the form of multiplicative image updating, which intrinsically enforced image non-negativity. The unique feature of this family was that the scaling factor was formed by the forward projection of the reconstructed image at the previous iteration, which is a unique feature in the "E-step" in an EM algorithm. Each

---

* Correspondence: larry.zeng@uvu.edu
[1]Department of Engineering, Utah Valley University, 800 W University Parkway, Orem, UT 84058, USA
[2]Department of Radiology and Imaging Sciences, University of Utah, 729 Arapeen Drive, Salt Lake City, UT 84108, USA
Full list of author information is available at the end of the article

member of the family had its own noise model. This work will extend this family of emission-EM-lookalike algorithms to Bayesian algorithms, by introducing a new factor. The three main features of the proposed algorithms comprise multiplicative updating with a non-negativity constraint, weighting by a projection noise model, and the incorporation of Bayesian constraints.

Many MAP algorithms in image reconstruction, especially in transmission tomography, employ the POCS methodology, which is an alternating optimization method. This breaks the objective function into different parts and optimizes each part separately. Our proposed method optimizes the objection function with all constraints considered simultaneously.

Methods

Modification of iterative Green's OSL algorithm

We first provide a brief review of Green's algorithm, before extending it. The iterative Green's OSL algorithm can be expressed as [1, 2].

$$x_{i,j}^{(n+1)} = \frac{x_{i,j}^{(n)}}{\sum_k a_{(i,j)k} + \beta U_{i,j}^{(n)}} \sum_k a_{(i,j)k} \frac{p_k}{\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}},$$

$$(1)$$

where $x_{i,j}^{(n)}$ is the reconstructed image pixel (i, j) at the nth iteration, p_k is the kth ray-sum measurement, $a_{(i,j)k}$ is the contribution of the pixel $x_{i,j}$ to the measurement p_k, β is a control parameter, and $U_{i,j}^{(n)}$ is the derivative of a penalty function V with respect to the image pixel $x_{i,j}^{(n)}$ at the nth iteration, i.e.,

$$U_{i,j}^{(n)} = \frac{\partial V}{\partial x_{(i,j)}^{(n)}}.$$

$$(2)$$

Using the approximation $1/(1+x) \approx 1-x$ when $|x| \ll 1$, the first factor $x_{i,j}^{(n)}/[\sum_k a_{(i,j)k} + \beta U_{i,j}^{(n)}]$ in algorithm (1) can be approximated as $\frac{x_{i,j}^{(n)}}{\sum_k a_{(i,j)k}}(1 - \frac{\beta}{\sum_k a_{(i,j)k}} U_{i,j}^{(n)})$.

Here, $\sum_k a_{(i,j)k}$, is in general not a constant. If $\beta/\sum_k a_{(i,j)k}$ is not a constant, then the constraint U is not uniformly enforced throughout the image. To improve the algorithm, we simply discard $\sum_k a_{(i,j)k}$ in $\beta/\sum_k a_{(i,j)k}$. Thus, our proposed modification of Green's algorithm is

$$x_{i,j}^{(n+1)} = \left(1 - \beta U_{i,j}^{(n)}\right) \frac{x_{i,j}^{(n)}}{\sum_k a_{(i,j)k}} \sum_k a_{(i,j)k} \frac{p_k}{\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}}.$$

$$(3)$$

We will gain further insight into this modification by rewriting both the original Green's algorithm (1) and the modified algorithm (3) in the additive-update form (that is, in the form of gradient descent). The additive form can be expressed as

$$x_{i,j}^{(n+1)} = x_{i,j}^{(n)} - \lambda_1 U_{i,j}^{(n)} - \lambda_2 \sum_k a_{(i,j)k} w_k \left(\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)} - p_k \right),$$

$$(4)$$

where

$$w_k = \frac{1}{\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}}$$

$$(5)$$

is the noise-weighting factor for the Poisson noise model and

$$\lambda_2 = \frac{x_{i,j}^{(n)}}{\sum_k a_{(i,j)k}}$$

$$(6)$$

is the step size for projection data fidelity minimization. In algorithm (4), λ_1 is the step size for Bayesian constraint minimization. For the original Green's algorithm (1),

$$\lambda_1^{(original)} = \beta x_{i,j}^{(n+1)} / \sum_k a_{(i,j)k},$$

$$(7)$$

whereas for the revised algorithm (3),

$$\lambda_1^{(revised)} = \beta \frac{x_{i,j}^{(n)}}{\sum_k a_{(i,j)k}} \sum_k a_{(i,j)k} \frac{p_k}{\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}}.$$

$$(8)$$

The most significant difference between algorithms (7) and (8) is that the factor $\lambda_1^{(original)}$ in (7) depends on the future image $x_{i,j}^{(n+1)}$, while the factor $\lambda_1^{(revised)}$ in (8) depends only on the current image $x_{i,j}^{(n)}$.

It is required that the image $x_{i,j}$ is non-negative. It can be observed from algorithm (3) that if $\beta U_{i,j}^{(n)} > 1$, then the sign of $x_{i,j}$ will alternate. Therefore, a necessary condition for the image to be non-negative is $\beta U_{i,j}^{(n)} < 1$. This intrinsic non-negativity constraint is guaranteed by the requirement that $\beta U_{i,j}^{(n)} < 1$ if the initial image is

positive. This can be readily observed by noticing that every factor in algorithm (3) is non-negative.

One way to prevent this from occurring is to introduce a sigmoid function ϕ, and to replace $\beta U_{i,j}^{(n)}$ by $\phi(\beta U_{i,j}^{(n)})$. There are many ways to define a sigmoid function ϕ. For example, one may choose $\phi(x) = x/\sqrt{1+x^2}$.

In deriving the Green's algorithm using prior information [1], it is necessary to evaluate the derivative of the energy function V, which carries the prior information. This energy function is defined for the updated image, which is not yet available. In Green's algorithm, an approximation is performed to evaluate this derivative of the energy function, using the current image to replace the future image. This approximation is termed "one-step-late".

The derivation of the EM-lookalike algorithms in ref. [10] was based on the noise variance model, unlike the conventional approach based on a random variable distribution function. Our derivation only considered two items: (1) the noise variance in the projections and (2) the non-negativity constraint for the image.

The derivation of the MAP in this study can been considered as an approximation of Green's MAP algorithm using $1/(1+x) \approx 1-x$ when $|x| \ll 1$. The proposed algorithms are in the form of "(1-βU) × (EM-lookalike)." When $\beta = 0$, this form is exactly the EM-lookalike form. The factor (1-βU) is new in this work, to minimize a Bayesian function V whose gradient is the function U. By driving U to zero, the Bayesian function V is minimized. The additive form algorithm (4) reveals that the proposed algorithms minimize the objective function

$$F = V + \frac{1}{2}\sum_k w_k \left(\sum_{i,j} a_{(i,j)k} x_{i,j} - p_k\right)^2, \qquad (9)$$

where the functions U and V are related as

$$U_{i,j}^{(n)} = \frac{\partial V}{\partial x_{(i,j)}^{(n)}}. \qquad (10)$$

For a different noise model, we can simply change the noise weighting w_k as in ref. [10].

This study builds on ref. [10], by considering a new energy function V and forcing its gradient U to zero. This point can be intuitively appreciated from the additive form algorithm (4).

From algorithm (6), we observe that the ML-EM algorithm's step size λ_2 is scaled by the image pixel value $x_{i,j}^{(n)}$ at the nth iteration. As a result, brighter objects converge faster than darker objects.

From algorithm (5), we observe that the weighting factor w_k is the reciprocal of the estimated mean value of

the kth ray-sum at the nth iteration. Note that w_k will change with different noise models.

From algorithm (7), we observe that λ_1 depends on the image value of the next iteration. This feature is undesirable, because it may cause the algorithm diverge. This undesirable feature has been removed from the revised algorithm, as shown in algorithm (8), where λ_1 depends only on the current image value.

The parameter λ_2 is intrinsically determined by the conventional ML-EM algorithm. The parameter λ_1 is affected by the parameter β. For any penalty function V, the parameter β is chosen by trial-and-error. When in doubt, a smaller positive β value should be chosen.

If the true solution with $\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}} = p_k$ and $U_{i,j} = 0$ exits, then it is straightforward to verify that the true solution is a fixed point of the proposed algorithm (3). In fact, letting $\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)} = p_k$ and $U_{i,j}^{(n)} = 0$, the right-hand side of (3) becomes $x_{i,j}^{(n)}$.

Modified algorithm for no weighting

We now consider a hypothetical imaging system, where the noise in the measurements is identically distributed with the same variance. In this case, noise weighting should not be utilized in the image reconstruction algorithm. The ML-EM lookalike algorithm for this hypothetical case is given as [10].

$$x_{i,j}^{(n+1)} = x_{i,j}^{(n)} \frac{\sum_k a_{(i,j)k} p_k}{\sum_k a_{(i,j)k} \sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}}. \qquad (11)$$

Using our strategy of introducing a simple new factor $(1-\beta U_{i,j}^{(n)})$, the Bayesian algorithm associated with algorithm (11) is proposed as
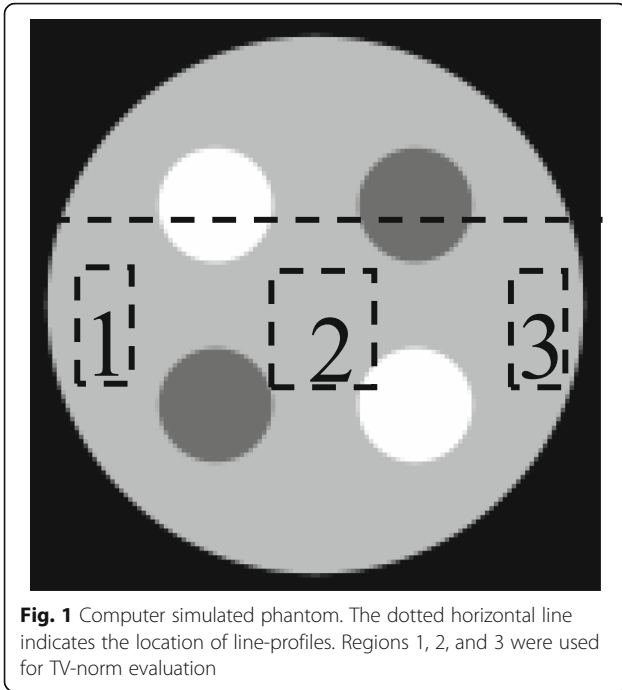
$$x_{i,j}^{(n+1)} = \left(1-\beta U_{i,j}^{(n)}\right) x_{i,j}^{(n)} \frac{\sum_k a_{(i,j)k} p_k}{\sum_k a_{(i,j)k} \sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}}. \qquad (12)$$

Modified algorithm for the transmission noise model

The variance of the transmission tomography sinogram is proportional to the exponential function of the sinogram's mean value [11]:

$$var(p_k) \propto \exp(\overline{p}_k) \approx \exp\left(\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}\right). \qquad (13)$$

An ML-EM lookalike algorithm for the transmission data is derived in ref. [10] as

Fig. 1 Computer simulated phantom. The dotted horizontal line indicates the location of line-profiles. Regions 1, 2, and 3 were used for TV-norm evaluation

$$x_{i,j}^{(n+1)} = x_{i,j}^{(n)} \frac{\sum_k \left[a_{(i,j)k} p_k \, \exp\left(-\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}\right) \right]}{\sum_k \left[a_{(i,j)k} \left(\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)} \right) \exp\left(-\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}\right) \right]}.$$

$$(14)$$

It is straightforward to modify algorithm (14) to a Bayesian algorithm, by introducing a new factor $\left(1-\beta U_{i,j}^{(n)}\right)$ as follows:

$$x_{i,j}^{(n+1)} = \left(1-\beta U_{i,j}^{(n)}\right) x_{i,j}^{(n)}$$
$$\times \frac{\sum_k \left[a_{(i,j)k} p_k \, \exp\left(-\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}\right) \right]}{\sum_k \left[a_{(i,j)k} \left(\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)} \right) \exp\left(-\sum_{\hat{i},\hat{j}} a_{(\hat{i},\hat{j})k} x_{\hat{i},\hat{j}}^{(n)}\right) \right]}.$$
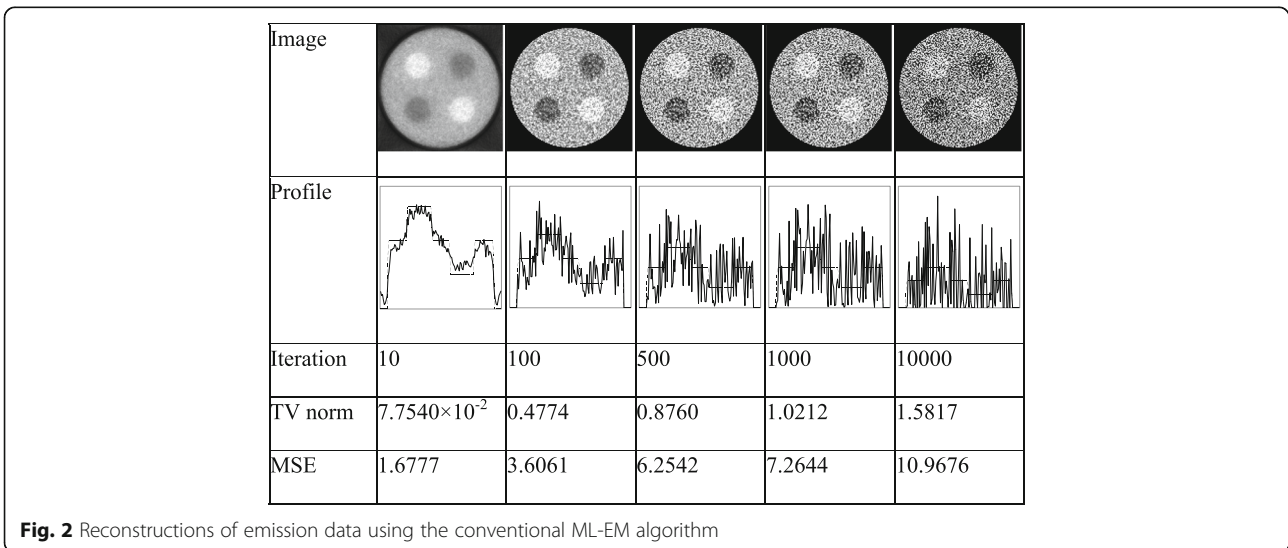
$$(15)$$

In general, a Bayesian algorithm can be readily obtained from a multiplicative-update image reconstruction algorithm by introducing a new factor $\left(1-\beta U_{i,j}^{(n)}\right)$. The resulting Bayesian algorithm remains multiplicative.

The TV penalty function

Any penalty function V can be employed in the proposed algorithm (3). Some constraints encourage smoothing, such as the maximum entropy constraint [18], because their main goal is denoising. Maximum entropy algorithms tend to over-smooth images, and as a result sharp edges are not maintained. Thus, maximum entropy algorithms are not popular for CT image reconstruction. On the other hand, TV-type constraints can reduce noise and maintain sharp edges when the parameters are suitably chosen. Here, we select the TV norm for a feasibility evaluation:

$$V = \sum_{i,j} \sqrt{\left(x_{i,j}-x_{i,j+1}\right)^2 + \left(x_{i,j}-x_{i+1,j}\right)^2}, \qquad (16)$$

where $x_{i,j}$ is a pixel value in a two-dimensional (2D) image. The associated derivative (2) is given as

Image					
Profile					
Iteration	10	100	500	1000	10000
TV norm	7.7540×10^{-2}	0.4774	0.8760	1.0212	1.5817
MSE	1.6777	3.6061	6.2542	7.2644	10.9676

Fig. 2 Reconstructions of emission data using the conventional ML-EM algorithm

	Image					
	Profile					
	Iteration	10	100	500	1000	10000
	TV norm	1.6973×10^{-2}	2.5174×10^{-2}	2.5771×10^{-2}	2.5772×10^{-2}	2.5772×10^{-2}
	MSE	1.5472	0.7708	0.8006	0.8028	0.8034

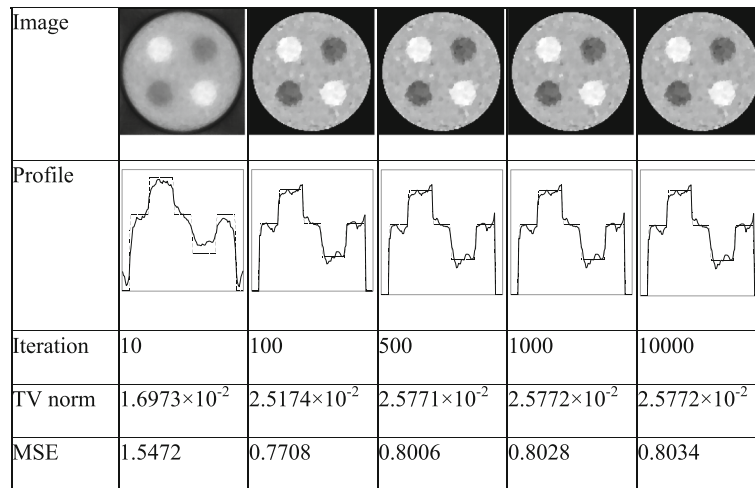**Fig. 3** Reconstructions of emission data using Green's OSL algorithm with $\beta = 1.2$

$$U_{i,j} \approx \frac{\left(x_{i,j} - x_{i,j+1}\right) + \left(x_{i,j} - x_{i+1,j}\right)}{\sqrt{\left(x_{i,j} - x_{i,j+1}\right)^2 + \left(x_{i,j} - x_{i+1,j}\right)^2 + \varepsilon}}$$
$$+ \frac{x_{i,j} - x_{i,j-1}}{\sqrt{\left(x_{i,j-1} - x_{i,j}\right)^2 + \left(x_{i,j-1} - x_{i+1,j-1}\right)^2 + \varepsilon}} \qquad (17)$$
$$+ \frac{x_{i,j} - x_{i-1,j}}{\sqrt{\left(x_{i-1,j} - x_{i-1,j+1}\right)^2 + \left(x_{i-1,j} - x_{i,j}\right)^2 + \varepsilon}}.$$

Here, the small value of ε is introduced to prevent the denominator being zero. In this study, $\varepsilon = 0.0001$ is adopted.

Computer simulations

Two sets of computer simulations were conducted, using emission and transmission noise models, respectively. The simulation setup for the emission data is as follows.

There were 180 projection views over 360°. The images were reconstructed in an array of size 128×128 (pixels). A parallel-hole collimation was assumed for the data generation. The detector had 128 detection bins, and the bin size was the same as the image pixel size.

A 2D circular phantom with a diameter of 120.32 pixels was employed in the simulations. The phantom, based on SPECT imaging, contained two small cold disks and two small hot disks, all with a diameter of 25.6 pixels, as shown in Fig. 1. The image intensity of the large circular disk was defined as 1 unit. The cold disks had an intensity value of 0.5, and the hot disks had an intensity value of 1.5. The projections were generated analytically, without using discrete pixels, and noisy projections were generated using the Poisson noise model. The total number of counts was approximately 2×10^6.
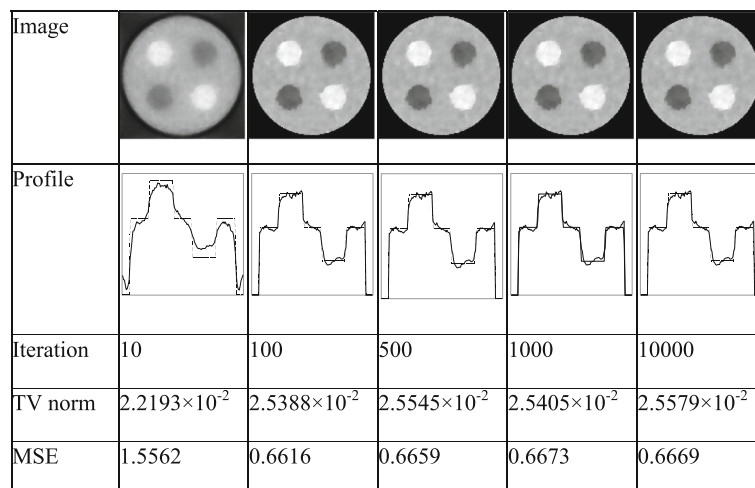

	Image					
	Profile					
	Iteration	10	100	500	1000	10000
	TV norm	2.2193×10^{-2}	2.5388×10^{-2}	2.5545×10^{-2}	2.5405×10^{-2}	2.5579×10^{-2}
	MSE	1.5562	0.6616	0.6659	0.6673	0.6669

Fig. 4 Reconstructions of emission data using the proposed algorithm with $\beta = 0.01$

Image					
Profile					
Iteration	10	100	500	1000	10000
TV norm	1.4310×10^{-4}	1.0318×10^{-3}	3.1736×10^{-3}	4.3850×10^{-3}	9.1151×10^{-3}
MSE	2.5019×10^{-2}	1.6944×10^{-2}	4.4573×10^{-2}	5.6655×10^{-2}	9.9884×10^{-2}
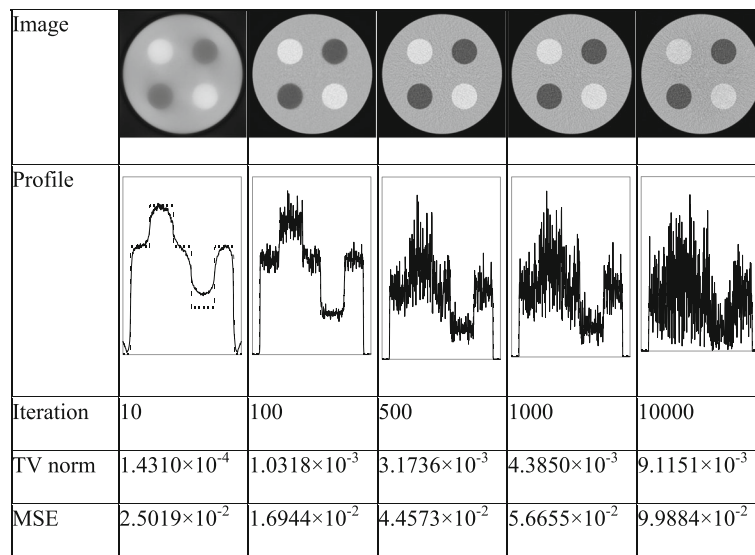
Fig. 5 Reconstructions of transmission data using algorithm (14) when $I_0 = 10{,}000$

The computer simulation setup for the transmission data was as follows. A parallel-beam imaging geometry was assumed. The image array was of size 512×512, the number of views was 400 over 180°, and the number of detection channels was 512. The transmission phantom looked similar to the emission phantom (Fig. 1), except four times larger. The pixel length was 0.5 mm. Furthermore, the attenuation coefficient was $0.0193 \, \mathrm{mm}^{-1}$ for the large disc, $0.0269 \, \mathrm{mm}^{-1}$ for the small circular bright regions, and $0.0083 \, \mathrm{mm}^{-1}$ for the small circular dark regions.

The transmission CT noise model was adopted for the sinogram data with very low counts, where the sinogram variance was proportional to the exponential function of the sinogram value. Two x-ray influxes were considered: $I_0 = 100$ and $I_0 = 10{,}000$.

Three regions were selected in the image for TV-norm noise evaluation. Note that the TV norm can measure the image fluctuation. These regions are depicted in Fig. 1. The average of the TV norms in these regions was employed as a figure-of-merit for noise evaluation. Furthermore, a line profile was provided for each
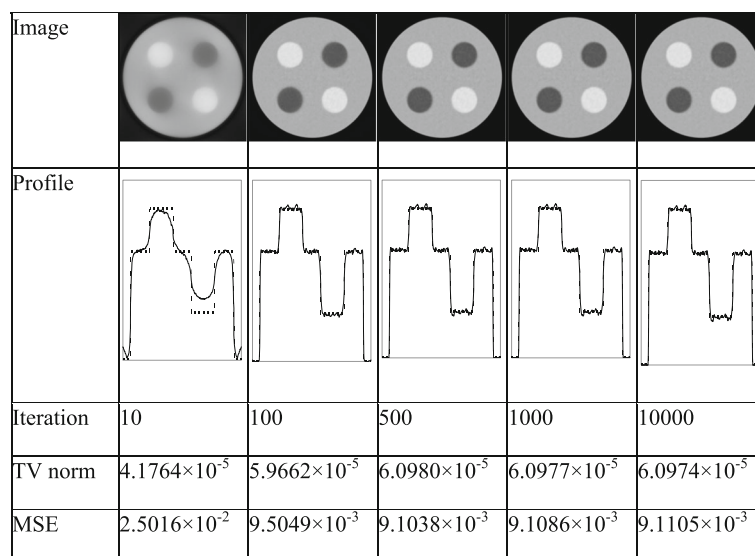
Image					
Profile					
Iteration	10	100	500	1000	10000
TV norm	4.1764×10^{-5}	5.9662×10^{-5}	6.0980×10^{-5}	6.0977×10^{-5}	6.0974×10^{-5}
MSE	2.5016×10^{-2}	9.5049×10^{-3}	9.1038×10^{-3}	9.1086×10^{-3}	9.1105×10^{-3}

Fig. 6 Reconstructions of transmission data using algorithm (15) with $\beta = 0.01$ when $I_0 = 10{,}000$

Image					
Profile					
Iteration	10	100	500	1000	10000
TV norm	1.1485×10^{-3}	9.6836×10^{-3}	2.0631×10^{-2}	2.2457×10^{-2}	2.5072×10^{-2}
MSE	3.1614×10^{-2}	1.6473×10^{-1}	3.5531×10^{-1}	3.8334×10^{-1}	4.3727×10^{-1}
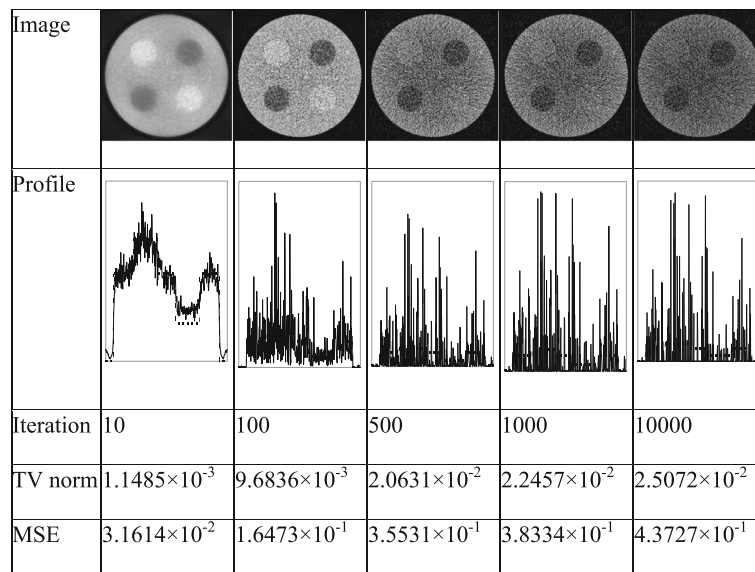
Fig. 7 Reconstructions of transmission data using algorithm (14) when $I_0 = 100$

reconstructed image. The location of the line profile is indicated in Fig. 1. As an additional figure-of-merit, the mean-squared-error (MSE) was also calculated between the reconstruction and true profiles, and this is reported in the figures.

Results

Emission data simulation results

Three algorithms were used to reconstruct the images: the conventional ML-EM algorithm (by setting $\beta = 0$ in either (1) or (9)), Green's OSL algorithm (1), and the proposed algorithm (9). The results are depicted in Figs. 2, 3, and 4, respectively, for the three algorithms. The proposed algorithm and Green's OSL algorithm yield similar performances.

The parameter β in the revised algorithm (3) is approximately equal to β in the original Green's algorithm (1) divided by the backprojection value of the constant 1. Roughly speaking, the β value in the original Green's algorithm is the β value in the revised algorithm times the number of view angles. In our example, $\beta = 1.2$ for the original Green's algorithm and $\beta = 0.01$ for the revised algorithm, and the number of view angles is 180. Thus, the regularization in Fig. 4 is a little stronger than that in Fig. 3.
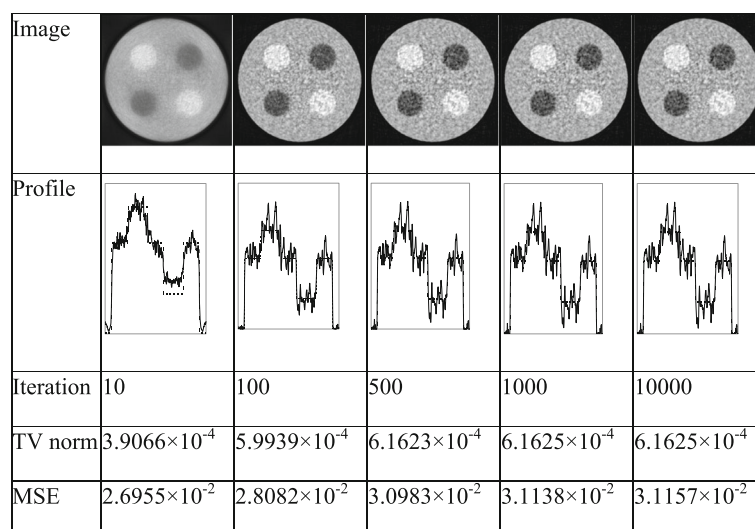

Image					
Profile					
Iteration	10	100	500	1000	10000
TV norm	3.9066×10^{-4}	5.9939×10^{-4}	6.1623×10^{-4}	6.1625×10^{-4}	6.1625×10^{-4}
MSE	2.6955×10^{-2}	2.8082×10^{-2}	3.0983×10^{-2}	3.1138×10^{-2}	3.1157×10^{-2}

Fig. 8 Reconstructions of transmission data using algorithm (15) with $\beta = 0.01$ when $I_0 = 100$

Image					
Profile					
Iteration	10	100	500	1000	10000
TV norm	1.2531×10^{-4}	3.6549×10^{-4}	7.5610×10^{-4}	8.0977×10^{-4}	8.1442×10^{-4}
MSE	5.6035×10^{-2}	1.6776×10^{-2}	1.3777×10^{-2}	1.4056×10^{-2}	1.4084×10^{-2}
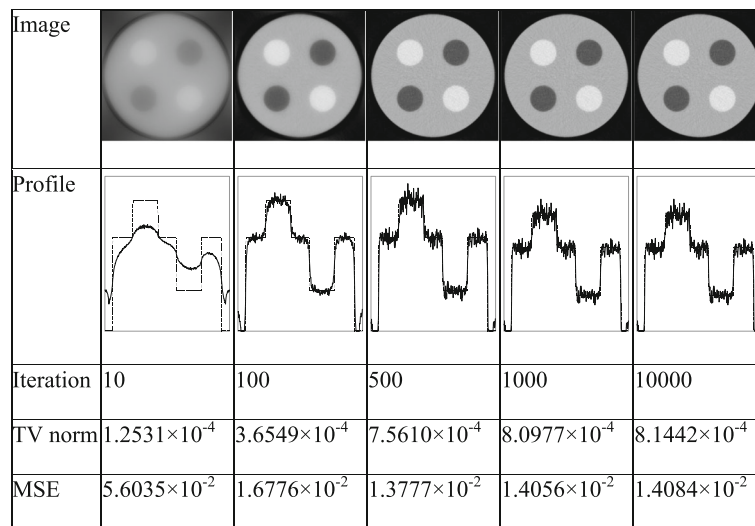
Fig. 9 Reconstructions with transmission data using the POCS algorithm in ref. [12] when $I_0 = 10,000$

Transmission data simulation results

Two algorithms were used to reconstruct the images: the EM-lookalike transmission algorithm (14) and proposed algorithm (15). For each algorithm, images were reconstructed with two noise levels. The results are presented in Figs. 5, 6, 7, and 8, for the two algorithms and two noise levels.

Finally, for comparison purposes we implemented the POCS algorithm proposed in ref. [12] and used it to reconstruct the transmission images. The results are presented in Figs. 9 and 10, for the lower and higher noise cases, respectively. We observe that our proposed simultaneous optimization algorithm performs better than the POCS algorithm proposed in ref. [12] in this task, in terms of the TV norm and MSE results.

It can be observed that the central region of the phantom appears darker in the Fig. 7. We hypothesize that noise may affect the convergence rate in an iterative algorithm. If a system of linear equations is more consistent, then the convergence rate may be faster. If the data is noisier and the system is less consistent, then the convergence rate may be slower.

We point out that when large 512×512 images are displayed as small binned-down images, as in Figs. 5–10, image details are lost. At iteration 10,000, all algorithms are considered converged. We zoom in on the upper-
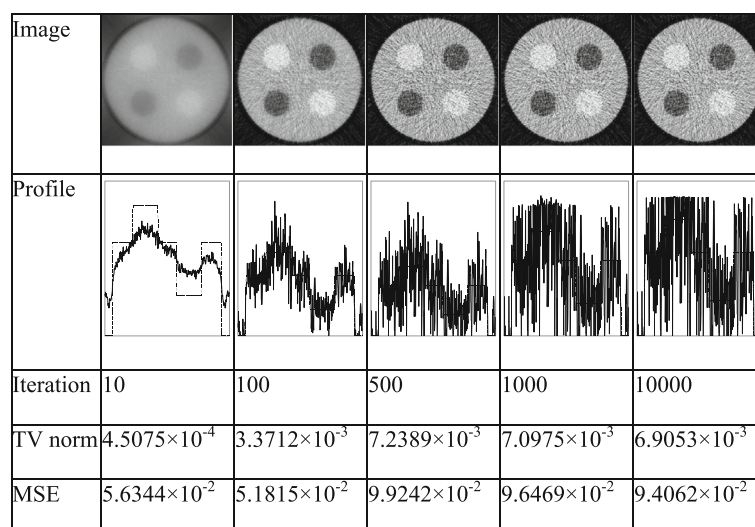

Image					
Profile					
Iteration	10	100	500	1000	10000
TV norm	4.5075×10^{-4}	3.3712×10^{-3}	7.2389×10^{-3}	7.0975×10^{-3}	6.9053×10^{-3}
MSE	5.6344×10^{-2}	5.1815×10^{-2}	9.9242×10^{-2}	9.6469×10^{-2}	9.4062×10^{-2}

Fig. 10 Reconstructions with transmission data using the POCS algorithm in ref. [12] when $I_0 = 100$
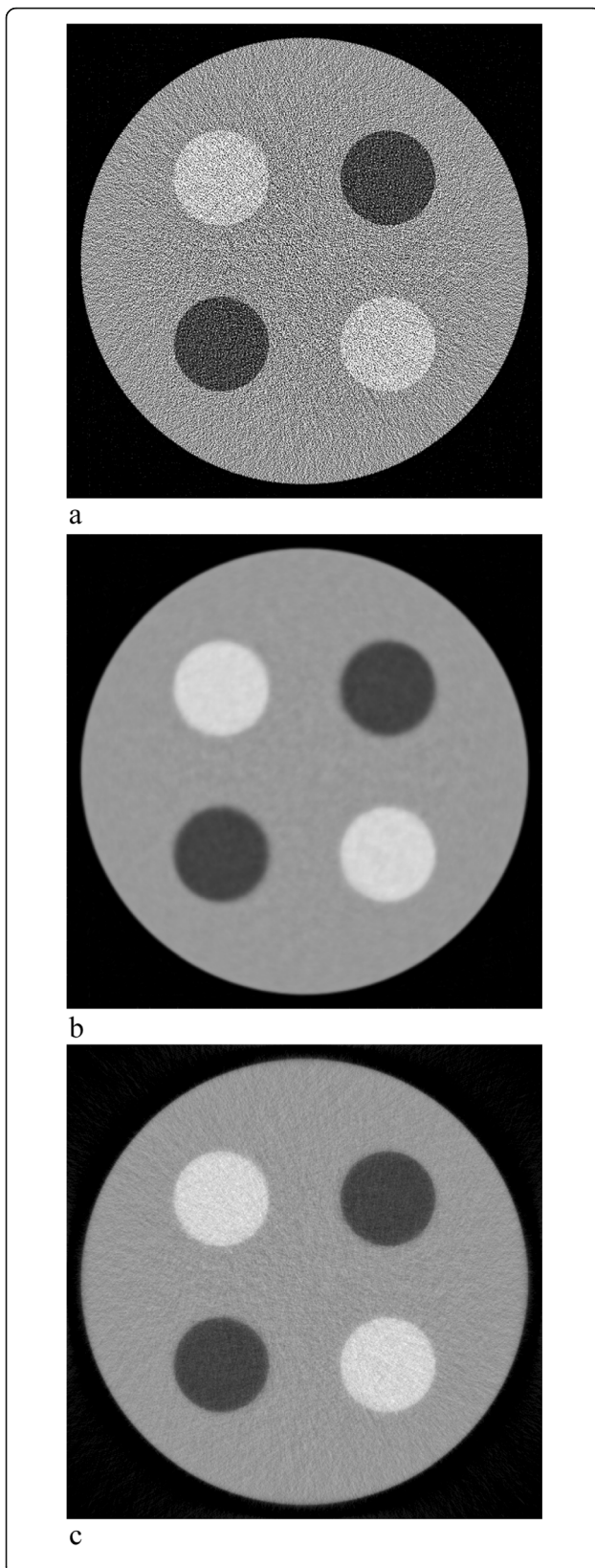
Fig. 11 Larger image display for lower noise transmission data reconstructions. **a**: View of the upper-right image in Fig. 5 (EM-lookalike algorithm); **b**: View of the upper-right image in Fig. 6 (proposed algorithm); **c**: View of the upper-right image in Fig. 9 (POCS algorithm in ref. [12])

right images in Figs. 5, 6, and 9 in Fig. 11. Here, one can better observe the differences between them. It is observed that the proposed Bayesian algorithms are effective in noise regularization, and stable as the iteration number increases.

The iterative POCS algorithm in our patient study provides better (yet noisier) spatial resolution than the proposed algorithm. The spatial resolution of an image reconstructed by the proposed iterative algorithms depends on the iteration number as well as the Bayesian penalty function. Usually, a larger iteration number gives a better spatial resolution, but a noisier reconstruction. The tradeoff between the spatial resolution and image noise is a main decision factor in selecting the iteration number. Suitable selection of the Bayesian penalty function, i.e., the constraints, plays an important role in the quality of the final reconstruction.

Conclusions

Our proposed algorithms are inspired by Green's OSL EM algorithm. The main novelty of this study is to propose a general methodology that extends EM-lookalike algorithms into MAP algorithms through a new multiplication factor $(1\text{-}\beta U)$. We claim that our approach can be extended to any multiplicative updating reconstruction algorithm, where image non-negativity is built in. Thus, the proposed algorithms also have an intrinsic non-negativity constraint. The proposed algorithms are simple to implement, and they simultaneously optimize all constraints (instead of using POCS).

We implemented the POCS algorithm presented in ref. [12] for transmission tomography, and we utilized the TV norm and MSE to evaluate the reconstructions. We observed that our proposed simultaneous optimization algorithm outperforms the POCS algorithm proposed in ref. [12] for our experiments.

Authors' contributions
All authors read and approved the final manuscript.

Authors' information
Gengsheng Zeng is with Department of Engineering, Utah Valley University and Department of Radiology and Imaging Sciences. Ya Li is with Department of Mathematics, Utah Valley University.

Availability of data and materials
Not applicable.

Competing interests
None of the authors have any competing interests in the manuscript.

Author details
[1]Department of Engineering, Utah Valley University, 800 W University Parkway, Orem, UT 84058, USA. [2]Department of Radiology and Imaging Sciences, University of Utah, 729 Arapeen Drive, Salt Lake City, UT 84108, USA. [3]Department of Mathematics, Utah Valley University, 800 W University Parkway, Orem, UT 84058, USA.

References
1. Green PJ (1990) On use of the EM algorithm for penalized likelihood estimation. J Roy Stat Soc: Ser B 52(3):443–452. https://doi.org/10.1111/j.2517-6161.1990.tb01798.x
2. Green PJ (1990) Bayesian reconstructions from emission tomography data using a modified EM algorithm. IEEE Trans Med Imag 9(1):84–93. https://doi.org/10.1109/42.52985
3. Panin VY, Zeng GL, Gullberg GT (1999) Total variation regulated EM algorithm [SPECT reconstruction]. IEEE Trans Nucl Sci 46(6):2202–2210. https://doi.org/10.1109/23.819305
4. Ellis S, Reader AJ (2017) Longitudinal multi-dataset PET image reconstruction. In: abstracts of 2017 IEEE nuclear science symposium and medical imaging conference, IEEE, Atlanta, GA, USA, 21-28 October 2017. https://doi.org/10.1109/NSSMIC.2017.8532657
5. Berker Y, Schulz V, Karp JS (2016) Discrete iterative algorithms for scatter-to-attenuation reconstruction in PET. In: abstracts of 2016 IEEE nuclear science symposium, medical imaging conference and room-temperature semiconductor detector workshop, IEEE, Strasbourg, France, 29 October-6 November 2016. https://doi.org/10.1109/NSSMIC.2016.8069455
6. Mair BA, Zahnen J (2006) A generalization of Green's one-step-late algorithm for penalized ML reconstruction of PET images. In: abstracts of 2006 IEEE nuclear science symposium conference record, IEEE, San Diego, CA, USA, 29 October-1 November 2006. https://doi.org/10.1109/NSSMIC.2006.356454
7. Zahnen JA (2006) Penalized maximum likelihood reconstruction methods for emission tomography. Dissertation, University of Florida. http://etd.fcla.edu/UF/UFE0015871/zahnen_j.pdf
8. Choi K, Lanterman AD (2007) Phase retrieval from noisy data based on minimization of penalized *I*-divergence. J Opt Soc Am A 24(1):34–49. https://doi.org/10.1364/JOSAA.24.000034
9. Fessler JA, Hero AO (1995) Penalized maximum-likelihood image reconstruction using space-alternating generalized EM algorithms. IEEE Trans Imag Proc 4(10):1417–1429. https://doi.org/10.1109/83.465106
10. Zeng GL (2018) Technical note: emission expectation-maximization look-alike algorithms for x-ray CT and other applications. Med Phys 45(8):3721–3727. https://doi.org/10.1002/mp.13077
11. Zeng GL, Wang WL (2016) Noise weighting with an exponent for transmission CT. Biomed Phys Eng Exp 2(4):045004. https://doi.org/10.1088/2057-1976/2/4/045004
12. Sidky EY, Pan XC (2008) Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. Phys Med Biol 53(17):4777–4807. https://doi.org/10.1088/0031-9155/53/17/021
13. Yan M, Chen JW, Vese LA, Villasenor J, Bui A, Cong J (2011) EM+TV based reconstruction for cone-beam ct with reduced radiation. In: Bebis G, Boyle R, Parvin B, Koracin D, Wang S, Kyungnam K et al (eds) Advances in visual computing. Springer, Berlin, Heidelberg, pp 1–10. https://doi.org/10.1007/978-3-642-24028-7_1
14. Hu ZL, Gao J, Zhang N, Yang YF, Liu X, Zheng HR, Liang D (2017) An improved statistical iterative algorithm for sparse-view and limited-angle CT image reconstruction. Sci Rep 7(1):10747. https://doi.org/10.1038/s41598-017-11222-z
15. De Pierro AR (1995) A modified expectation maximization algorithm for penalized likelihood estimation in emission tomography. IEEE Trans Med Imag 14(1):132–137. https://doi.org/10.1109/42.370409
16. Erdogan H, Fessler JA (1999) Monotonic algorithms for transmission tomography. IEEE Trans Med Imag 18(9):801–814. https://doi.org/10.1109/42.802758
17. Erdogan H, Fessler JA (1999) Ordered subsets algorithms for transmission tomography. Phys Med Biol 44(11):2835–2851. https://doi.org/10.1088/0031-9155/44/11/311
18. Elfving T (1989) An algorithm for maximum entropy image reconstruction from noisy data. Math Comput Model 12(6):729–745. https://doi.org/10.1016/0895-7177(89)90358-0

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