Published in final edited form as:

IEEE Trans Med Imaging. 2011 April ; 30(4): 946–956. doi:10.1109/TMI.2010.2098036.

Reconstruction of Emission Tomography Data Using Origin Ensembles

Arkadiusz Sitek [Member, IEEE]

Department of Radiology, Harvard Medical School and Brigham and Women's Hospital, Boston, MA, 02115 USA

Arkadiusz Sitek: asitek@bwh.harvard.edu

Abstract

A new statistical reconstruction method based on origin ensembles (OE) for emission tomography (ET) is examined. Using a probability density function (pdf) derived from first principles, an ensemble expectation of numbers of detected event origins per voxel is determined. These numbers divided by sensitivities of voxels and acquisition time provide OE estimates of the voxel activities. The OE expectations are shown to be the same as expectations calculated using the complete-data space. The properties of the OE estimate are examined. It is shown that OE estimate approximates maximum likelihood (ML) estimate for conditions usually achieved in practical applications in emission tomography. Three numerical experiments with increasing complexity are used to validate theoretical findings and demonstrate similarities of ML and OE estimates. Recommendations for achieving improved accuracy and speed of OE reconstructions are provided.

Index Terms

Tomographic Reconstruction; Statistical analysis; Emission Tomography; SPECT; PET; Markov Chains

I. Introduction

In emission tomography, volumetric images are estimated from ET projection data using various reconstruction algorithms. In this paper, a tomographic reconstruction algorithm based on principle of origin ensembles is examined.

There are two general approaches to image reconstruction in ET, namely analytic and iterative approaches. In current clinical applications the iterative reconstruction methods [1], [2] are most frequently used because they proved to be more robust to noise compared to analytic approaches [3]. The most popular iterative approaches are based on the principle of the maximum likelihood (ML) where the likelihood of observed data is maximized as function of voxel activities. The expectation maximization (EM) [4], [5], [6] algorithm or direct maximization using for example conjugate gradient can be used for finding the ML estimates. The ML estimate has a very high variance and it is not useful for clinical purposes. Usually some a priori knowledge is used to regularize (reduce variance) of a ML-based method. The most common approach to regularization is to enforce neighboring voxels to have similar estimates of activity, but other approaches can also be used. For

review on these methods see [1], [2]. A serious drawback of the iterative methods is that they require so called projection and backprojection operations. In the projection operation, the expected data based on a current image estimate is calculated. In the backprojection operation the image estimate is modified based on ratios or differences between acquired data and expected data obtained in the projection operation. Both of these operations are compute intensive and using them in modeling of complex acquisition geometries is impractical (see the discussion for particular examples of such systems). This prompted a search for an algorithm which would be robust to noise similarly to the statistical iterative methods and at the same time would not require projection/backprojection operations. Such algorithm was presented in [7] which was based on a calculation of the expected number of detected events emitted per voxel given the data. The convergence of the algorithm to the steady state was shown in [7]. This algorithm is referred to as the OE algorithm.

It will be shown in this paper the OE and ML estimates for typical inverse problems found in ET are very similar (noise artifacts are the same) and differences between the results of these two methods are likely to be negligible for practical applications. It will also be shown that the OE estimate agrees better with the ML estimate with increasing number of detected events and fewer zero activity voxels. When number of detected events approaches the infinity OE and ML estimates become identical.

The main purpose of this work is to demonstrate that the result of the OE algorithm approximates ML estimate. This will be done by placing the OE in the context of statistical reconstruction theory in section II-C. Next, in section II-D, two approximations to the OE estimate are used under which the approximated-OE estimate and the ML estimate are identical. These two approximations hold well in practical applications of ET as it is demonstrated in next sections. In section III a thought experiment is used for which the OE estimate (ensemble expectation) can be obtained with perfect accuracy directly calculating OE expectations. In next two sections IV and V examples of the OE reconstructions (Monte Carlo (MC) integrations of the ensemble) and comparison with the ML estimates for 2D-SPECT and 3D-PET are provided.

II. Theory

A. Introduction to notation for inverse problems in ET

The image volume is divided into volume elements indexed by i which ranges from 1 to I . The value of I represents a total number of volume elements. In most cases voxels are used as the volume elements, but in general other elements can be utilized such as point clouds [8] or blobs [9]. For the case of overlapping elements such blobs or point clouds, the OE algorithm can also be implemented, however for each event location a stochastic decision needs to be made as to which blob or point “contains” the event. These consideration are irrelevant for the purpose of this paper, it is noteworthy mentioning however, that the OE approach can be applied to overlapping volume elements. Nonetheless, to simplify the concepts in the following considerations, volume elements will be referred to as voxels keeping in mind that other image representations can also be used. Projection elements will always be referred to as bins and elements of the reconstruction area as voxels. A number of events detected during the acquisition in a bin k is denoted by g_k where the index k enumerates projection bins and goes from 1 to K . The imaging system is described by a *system matrix* with elements a_{ki} . An element a_{ki} describes the probability that an event emitted in the volume element (voxel) i is detected in the bin k . Note, that for a detected event n in the projection bin k the a_{kn} is a probability that if a detected event n (n goes from 1 to N) was emitted in i it is detected in k . For a given n the bin in which event n was detected is k_n . For clarity of notation double indices are omitted in following sections and simply $a_{nki} = a_{ni}$. It is assumed that system matrix contains information about object

attenuation. Although effects as Compton scatter can also be included in the definition of the system matrix at this point, only attenuation is included since in the OE approach Compton scatter and randoms (for PET) effects can be handled more efficiently in a different manner than for other reconstruction methods and describing of these approaches in full details is beyond the scope of this work. See discussion section for more information on OE implementation of scatter correction and randoms for PET. The sensitivity of voxels is denoted by ϵ_i and by definition is equal to

$$\epsilon_i = \sum_{k=1}^K \alpha_{ki}. \quad (1)$$

In general $\hat{f}_i$ is an estimate obtained by any method. Based on an estimate $\hat{f}_1 \dots \hat{f}_I$ the expected number of counts in bin k calculated

$$\hat{g}_k = \sum_{i=1}^I \alpha_{ki} \hat{f}_i. \quad (2)$$

The above equation defines a projection operation. For list-mode data the projection operation for each detected event n is defined as

$$\hat{g}_n = \sum_{i=1}^I \alpha_{ni} \hat{f}_i. \quad (3)$$

The value of $\hat{g}_n$ indicates an estimate of number of counts in the projection bin in which event n was detected. The value of α_{ni} indicates an element of system matrix of the voxel i and the projection bin in which event n was detected.

B. Origin ensemble estimator

This section is a summary of the OE algorithm presented in details elsewhere [7]. Note that the OE algorithm defined as Markov Chains Monte Carlo estimation of ensemble expectation is rigorously derived and its convergence to equilibrium proven in [7].

In photon limited tomography, photons are detected that are the result of a nuclear process (event) that occurred in one of the voxels. Origin ensemble (OE) is created by assuming that the origin of the event that generated detected photon (or photons) could have happened in any of the I voxels with chance that is equal or higher than zero. N detected events and I voxels give I^N possibilities of assigning locations of origins of all events to the voxels. The specification of event origin locations for all N events defines a *state* s in the ensemble of all I^N states (origin ensemble). The number of event origins in voxel i for state s is denoted as c_{si} . The estimate of activity of a voxel i for a state s is defined as $\hat{f}_{si} = c_{si} / \epsilon_i$. In other words, the estimate of activity in a given voxel is the number of emitted and detected events divided by the sensitivity of the voxel. This value of activity maximizes Poisson probability that c_{si} events were emitted from this voxel.

The probability that the origin of event n is located in voxel i assuming activity estimate in this voxel $\hat{f}_{si}$:

$$\rho_{sni} \sim \alpha_{ni} \hat{f}_{si} = \alpha_{ni} c_{si} / \epsilon_i. \quad (4)$$

A state s is defined by locations of all events, therefore it follows using the Eq. 4 that the probability of a state s is expressed as:

$$\Pi(s) \sim \prod_{n=1}^N \alpha_{ni} c_{si} / \varepsilon_i = \prod_{n=1}^N \alpha_{ni} \prod_{i=1}^I c_{si} / \varepsilon_i \quad (5)$$

Note that in the above equation the index i indicates the voxel in which event n is located for state s . When computing the product $\prod_{n=1}^N c_{si} / \varepsilon_i$ over the list of events, each voxel will be 'visited' c_{si} times, once for each of the c_{si} events contained in the voxel, leading to a contribution of $(c_{si} / \varepsilon_i)^{c_{si}}$ for this particular voxel. This means that the product over the event list, n , of c_{si} / ε_i can be converted into a product over voxels, j , of $(c_{sj} / \varepsilon_j)^{c_{sj}}$, to obtain:

$$\Pi(s) \sim \prod_{n=1}^N \alpha_{ni} \prod_{j=1}^I (c_{sj} / \varepsilon_j)^{c_{sj}} \quad (6)$$

The probability of a state is defined only in terms of a proportionality relation in Eq. 6. In general, the normalization constant can be found by summing Eq. 6 over all states. However, the knowledge of the normalization constant is not needed for the OE estimator as it will soon be demonstrated.

Considering two states from the ensemble, state s and state s' that differ by a location of a single origin of an event n , without losing generality it can be assumed that in the state s , the event n is located in i and in state s' the event n is located in i' . Using Eq. 6 the following is obtained (most of the terms in the product in Eq. 6 are canceled when ratio is considered):

$$\frac{\Pi(s')}{\Pi(s)} = \frac{\alpha_{ni'} (c_{si} - 1)^{c_{si}-1} (c_{si'} + 1)^{c_{si'}+1} / \varepsilon_{i'}}{\alpha_{ni} (c_{si})^{c_{si}} (c_{si'})^{c_{si'}} / \varepsilon_i}, \quad (7)$$

Using the above Eq. 7, the OE reconstruction algorithm that determines average number of event origins per voxel by calculation of the ensemble expectation is defined using the following Markov Chains Monte Carlo procedure:

- Step 1** Create starting state s_0 by randomly selecting possible origin locations of all detected events in voxels in the reconstruction area. Origins of all detected events indexed by n are placed in voxels i such that $\alpha_{ni} > 0$. This guarantees that $\Pi(s_0) > 0$. The final OE estimate will not depend on the starting point, so the algorithm used for the creation of the starting point does not affect the final result.
- Step 2** Randomly select an event n and note voxel i in which the origin of n is located.
- Step 3** Randomly select a new voxel i' that would contain the origin of n if the move (Step 4) is successful.
- Step 4** Stochastically move the event n to the new voxel with the transition probability equal to

$$p_{n(i \rightarrow i')}(c_{si}, c_{si'}) = \min \left(1, \frac{\Pi(s')}{\Pi(s)} \right) = \min \left(1, \frac{\alpha_{ni'}(c_{si} - 1)^{c_{si}-1} (c_{si'} + 1)^{c_{si'}+1} / \epsilon_{i'}}{\alpha_{ni}(c_{si})^{c_{si}} (c_{si'})^{c_{si'}} / \epsilon_i} \right), \quad (8)$$

Step 5 Go to Step 2.

Steps 2 through 5 repeated N times constitute one iteration of the OE algorithm. The form of the transition probability Eq. 9 follows the work of Metropolis *et al.* [10]. It guarantees that the algorithm reaches equilibrium (stationary state) since $p_{n(i \rightarrow i')}(c_{si}, c_{si'})$ expressed by Eq. 9 and $p_{n(i' \rightarrow i)}(c_{si'}, c_{si}) = \min \left(1, \frac{\Pi(s)}{\Pi(s')} \right)$ satisfy detailed balance condition [7], [10], [11] given below:

$$\frac{\Pi(s')}{\Pi(s)} = \frac{p_{n(i \rightarrow i')}(c_{si}, c_{si'})}{p_{n(i' \rightarrow i)}(c_{si'}, c_{si})} \quad (9)$$

where states s and s' differ by a location of single origin of the event n . For the state s origin of n is in the voxel i and for the state s' in the voxel i' . Once in the equilibrium, the above algorithm will generate states with frequencies proportional to probabilities of these states defined by Eq. 6 therefore effectively calculating ensemble expectation using Monte Carlo approach. By measuring c_{si} for these equilibrium states (ideally over a very large number of equilibrium states) and averaging these values, the ensemble average (ensemble expectation) c_i^* is determined. All values corresponding to OE averages (ensemble expectations) are indicated with the superscript $*$.

Finally, the OE estimator is defined as:

$$\widehat{f}_i^{OE} = c_i^* / \epsilon_i. \quad (10)$$

Note that the estimator is defined based on the ensemble expectation and not based on a state from the ensemble with the maximum probability.

C. Origin ensemble in the context of statistical reconstruction

In this section, it is demonstrated that the expectations calculated using the OE algorithm are equal to the expectations of maximized activities calculated over the complete-data space using Poisson likelihoods.

First, the concept of the complete-data space is recalled [5], [12]. The (unobservable) complete-data space $\mathbf{h}$ is described by h_{ki} which indicate numbers of events emitted in voxels i and detected in projection elements k . There are a number of origin ensemble states that correspond to the same complete-data state. For example, a single data value from the complete-space $h_{ki} = 2$ with $g_k = 5$ would correspond to $\frac{5!}{2!(5-2)!} = 10$ different equal probability-states from origin ensemble. In general, it can be shown that for any complete-

data state described by a vector $\mathbf{h}$ there are $\prod_{k=1}^K g_k! / \prod_{k,i=1}^{K,I} h_{ki}!$ equal-probability origin ensemble states.

In order to show that expectations are the same, it is suffice to show that the Poisson

likelihood of complete-data $\mathcal{L}(\mathbf{h})$ is equal to $\Pi(s)$ multiplied by $\prod_{k=1}^K g_k! / \prod_{k,i=1}^{K,I} h_{ki}!$. The complete-data likelihood is evaluated for voxel activities ($\hat{f}_{si} = c_{si} / \epsilon_i$) which maximize

probability that voxels emitted c_{si} events. It follows that the Poisson likelihood of complete-data state for maximized activities is from definition:

$$l(\mathbf{h}) = \frac{\prod_{k,i=1}^{K,I} e^{-\alpha_{ki} c_{si} / \varepsilon_i} (\alpha_{ki} c_{si} / \varepsilon_i)^{h_{ki}}}{\prod_{k,i=1}^{K,I} h_{ki}!}. \quad (11)$$

Since the $\sum_{k=1}^K \alpha_{ki} = \varepsilon_i$ (Eq. 1) it follows that a term from $l(\mathbf{h})$ is evaluated as follows:

$$\prod_{k,i=1}^{K,I} e^{-\alpha_{ki} c_{si} / \varepsilon_i} = \prod_{i=1}^I e^{-\varepsilon_i c_{si} / \varepsilon_i} = \prod_{i=1}^I e^{-c_{si}} = e^{-N} \quad (12)$$

Using the above Eq. 12 and $\sum_{k=1}^K h_{ki} = c_{si}$, the likelihood $l(\mathbf{h})$ transforms to:

$$l(\mathbf{h}) = \frac{e^{-N} \prod_{k,i=1}^{K,I} (\alpha_{ki})^{h_{ki}} \prod_{i=1}^I (c_{si} / \varepsilon_i)^{c_{si}}}{\prod_{k,i=1}^{K,I} h_{ki}!} \quad (13)$$

Defining constant $a = e^{-N} / \prod_{k=1}^K g_k!$ and rewriting the above equation for list-mode data:

$$l(\mathbf{h}) = a \frac{\prod_{k=1}^K g_k!}{\prod_{k,i=1}^{K,I} h_{ki}!} \prod_{n=1}^N \alpha_{ni} \prod_{i=1}^I (c_{si} / \varepsilon_i)^{c_{si}} = \frac{\prod_{k=1}^K g_k!}{\prod_{k,i=1}^{K,I} h_{ki}!} \Pi(s). \quad (14)$$

Although the above shows that effectively the complete-data expectation is obtained by the OE algorithm, it is far from obvious that expectation of maximized activities in complete-data space is similar to voxel activities that give the maximum likelihood of the incomplete (observed) data.

D. Approximation of OE Estimate

In this section, it is shown that under certain approximations (expressed by Eqs. 15 and 17), the OE estimate is the same as the ML estimate for the ET inverse problems. All considerations that will follow in this section apply to the OE that is in equilibrium and are general for any linear tomographic system with photon limited data.

In the stationary state of the OE algorithm, the $\pi_{ni}^* \geq 0$ is the equilibrium probability that origin of the event n is located in a voxel i .

The following approximation is made first (see Appendix A

$$\frac{\pi_{ni}^*}{\pi_{ni}^*} \approx \frac{P_{n(i \rightarrow i')}^*}{P_{n(i' \rightarrow i)}^*} \approx \frac{P_{n(i \rightarrow i')}(c_i^*, c_{i'}^*)}{P_{n(i' \rightarrow i)}(c_{i'}^*, c_i^*)}. \quad (15)$$

By $P_{n(i \rightarrow i')}^*$ the average transition probability from voxel i to voxel i' for the event n is indicated. The above approximation consists of two parts. First coming from

($\pi_{ni}^* P_{n(i \rightarrow i')}^* \approx \pi_{ni'}^* P_{n(i' \rightarrow i)}^*$) states that the balance in “probability transfer” between any two

voxels for every event can be approximated by the balance of the *average* probability transfer between any two voxels for every event. The average probability transfer from voxels i to i' for event n is defined as the product of the average probability that the event n is in voxel i (π_i^*) multiplied by the average chance that it is moved to voxel i' . Obviously there is a balance in probability transfer for every state due to detailed balance equation (see appendix A for derivation). The second part of the approximation is that the average probability transfer can be approximated by the transition probability for average values of number of events in voxels. These approximations will hold very well if $P_{n(i \rightarrow i')}(c_i^*, c_{i'}^*)$ and $P_{n(i' \rightarrow i)}(c_{i'}^*, c_i^*)$ are approximately the same for states obtained when the algorithm is in equilibrium. In section III this approximation is studied in more details.

Using Eq. 9 in the ratio expressed by Eq. 15:

$$\frac{P_{n(i \rightarrow i')}(c_i^*, c_{i'}^*)}{P_{n(i' \rightarrow i)}(c_{i'}^*, c_i^*)} = \frac{\alpha_{ni'}(c_i^* - 1)^{c_i^* - 1}(c_{i'}^* + 1)^{c_{i'}^* + 1} / \varepsilon_{i'}}{\alpha_{ni}(c_i^*)^{c_i^*}(c_{i'}^*)^{c_{i'}^*} / \varepsilon_i} = \frac{\alpha_{ni'} c_{i'}^* / \varepsilon_{i'}}{\alpha_{ni} c_i^* / \varepsilon_i} \left[\left(1 + \frac{1}{(c_i^* - 1)} \right)^{(c_i^* - 1)} \right]^{-1} \left[\left(1 - \frac{1}{(c_{i'}^* + 1)} \right)^{(c_{i'}^* + 1)} \right]^{-1} \quad (16)$$

Now using the formula that $(1 \pm 1/z)^z \approx e^{\pm 1}$ the following is obtained:

$$\left[\left(1 + 1/(c_i^* - 1) \right)^{(c_i^* - 1)} \right]^{-1} \left[\left(1 - 1/(c_{i'}^* + 1) \right)^{(c_{i'}^* + 1)} \right]^{-1} \approx 1. \quad (17)$$

Using Eqs. 15, 17, and 17:

$$\frac{\pi_{ni'}}{\pi_{ni}^*} \approx \frac{\alpha_{ni'} c_{i'}^* / \varepsilon_{i'}}{\alpha_{ni} c_i^* / \varepsilon_i}. \quad (18)$$

It is obvious that the approximation that was used above is better for higher c_i^* and $c_{i'}^*$. The end result of the above approximations expressed by the Eqs. 15 and 17 states that in the equilibrium, the ratio of probabilities of having event n in voxel i and i' is proportional to the product of the ratio of elements of the system matrix for this events and the ratio of OE estimated activities in these voxels $\widehat{f}_i^{OE} / \widehat{f}_{i'}^{OE}$ since $\widehat{f}_i^{OE} = c_i^* / \varepsilon_i$ and $\widehat{f}_{i'}^{OE} = c_{i'}^* / \varepsilon_{i'}$ from the definition in Eq 10.

In order to show that the OE estimator after applying the above approximation is equal to the ML estimator, the normalizing condition that applies to π 's is noted:

$$\sum_{i'=1}^I \pi_{ni'}^* = 1 \quad (19)$$

Substituting in Eq. 19 value of $\pi_{ni'}^*$ from Eq. 18 the following is obtained:

$$\pi_{ni}^* \sum_{i'=1}^I \frac{\alpha_{ni'} c_{i'}^* / \varepsilon_{i'}}{\alpha_{ni} c_i^* / \varepsilon_i} \approx 1, \quad (20)$$

which further transforms to:

$$\pi_{ni}^* \approx \frac{1}{\sum_{i'=1}^I \frac{\alpha_{ni'} c_{i'}^* / \varepsilon_{i'}}{\alpha_{ni} c_i^* / \varepsilon_i}} = \frac{\alpha_{ni} c_i^*}{\varepsilon_i} \frac{1}{\sum_{i'=1}^I \alpha_{ni'} c_{i'}^* / \varepsilon_{i'}}. \quad (21)$$

From the definition, the average number of events c_i^* in voxel i is a sum over π_{ni}^* , therefore

$$c_i^* = \sum_{n=1}^n \pi_{ni}^*. \quad (22)$$

Now using π_{ni}^* from Eq 21 the above becomes

$$c_i^* \approx \frac{c_i^*}{\varepsilon_i} \sum_{n=1}^N \frac{\alpha_{ni}}{\sum_{i'=1}^I \alpha_{ni'} c_{i'}^* / \varepsilon_{i'}}. \quad (23)$$

Taking into account the definition of OE estimator $\widehat{f}_i^{OE} = c_i^* / \varepsilon_i$ (Eq. 10) and noting that c_i^* 's cancel each other on both sides of the Eq. 23, the above transforms to

$$\sum_{n=1}^N \frac{\alpha_{ni}}{\sum_{i'=1}^I \alpha_{ni'} \widehat{f}_{i'}^{OE}} \approx \varepsilon_i. \quad (24)$$

On the other hand for the list-mode maximum likelihood approach, the ML estimate satisfies the following (see Appendix B for derivation)

$$\sum_{n=1}^N \frac{\alpha_{ni}}{\sum_{i'=1}^I \alpha_{ni'} \widehat{f}_{i'}^{ML}} = \varepsilon_i \text{ for } i=1 \dots I \quad (25)$$

The above equation is the same as Eq. 24. Since the likelihood function is strictly convex only one maximum value of likelihood exists so $\widehat{\mathbf{f}}^{ML}$ and approximated $\widehat{\mathbf{f}}^{OE}$ achieve the same maximum. For an inverse problem for which the null space is empty which indicates that $\widehat{\mathbf{f}}^{ML}$ the following is also true:

$$\widehat{f}_i^{OE} \approx \widehat{f}_i^{ML} \text{ for } i=1 \dots I. \quad (26)$$

III. Reconstruction of point source – thought Experiment

In the previous section, it was shown that the approximation of the OE estimate equals the ML estimate. In this section, a thought experiment will be used in which a point source is imaged using two tomographic projections in order to investigate the effects of the approximations on the OE estimate. Due to symmetries and simplicity of the experiment, the exact OE estimate was obtained by integrating over all states of the ensemble and solving a non-linear equation. The OE estimate after the applying approximation expressed by Eq. 15 (by not applying approximation expressed by 17) by solving a non-linear equation is also determined. Differences between these two estimates indicate the effect of that approximation described by Eq. 15 on the OE estimate. Obviously, the application of the

second approximation (Eq. 17) makes the OE estimate equal to the ML estimate, which is also demonstrated in this section.

A. Point source experiment setup

In the experiment, the reconstruction area is defined as the $(B + 1) \times (B + 1)$ pixel matrix in 2D. The center pixel emits $2N$ photons which all are detected in 2 projections (vertical and horizontal) separated by the 90° angle. In each projection, only one bin detects exactly N counts. It is assumed that all non-zero elements of the system matrix are the same and equal to $1/2$ and what follows sensitivities of all pixels are also the same and equal to 1. See Figure 1 for details of the setup. Note that the setup is vulnerable to weaknesses of the

approximation of $c_i^* \gg 1$ and $c_l^* \gg 1$ used in Eq. 17 since it is expected that pixels on vertical and horizontal lines will be reconstructed by the OE with a very low value of counts. Therefore it is expected that OE estimate may be quite different than the ML estimate. It is obvious that the ML estimate for such system corresponds to the image with the center pixel with activity $2N$ and all other pixels with 0.

B. Exact (not Monte Carlo) OE Estimate

The exact OE estimate is obtained by integrating all states in an ensemble to obtain accurate ensemble average. It is computationally possible to accomplish for a special case discussed in this section due to symmetries. For the OE algorithm in the equilibrium, three groups of pixels are identified due to symmetries. For first group consisting of the single center pixel, the probability that any event was emitted from the center pixel is the same for all $2N$ events and equal to π_x^* . For B pixels that are on vertical/horizontal lines (vertical/horizontal shaded pixels in Figure 1) the probability that events were emitted from them and detected at vertical/horizontal projections is equal to π_l^* and is the same for all N photons detected in vertical/horizontal projections. The third group consists of all other pixels for which $\pi_0^* = 0$. Note that π_x^* , π_l^* , and π_0^* are the same for all detected photons due to symmetry. Therefore, the index n is omitted.

Considering the vertical projection, for a given state s during the OE algorithm in equilibrium the number of events in the central pixel and detected in vertical projection is c_{sxv} . This number follows the binomial distribution $P_{\text{Bin}}(c_{sxv}, N, \pi_x^*)$. For the horizontal projection for which two pixels (central pixel x and any pixel l on shaded horizontal line) are considered the number of events in the central pixel c_{sxh} and pixel l c_{sl} will follow trinomial distribution with $P_{\text{Tri}}(c_{sxh}, c_{sl}, N, \pi_x^*, \pi_l^*)$ (see appendix C for mathematical formulas describing these distributions). For an event detected in the horizontal projection for a state s in the ensemble, the probability that the event is in pixel x is equal to c_{sxh}/N and in pixel l c_{sl}/N . Since the detailed balance equation guarantees that in equilibrium no net flow of probability from/to any pixel exists therefore:

$$\sum_{c_{sxv}=0}^N \sum_{c_{sxh}=0}^N \sum_{c_{sl}=0}^N P_{\text{Bin}}(c_{sxv}, N, \pi_x^*) P_{\text{Tri}}(c_{sxh}, c_{sl}, N, \pi_x^*, \pi_l^*) \left(\frac{c_{sxh}}{N} p_{x \rightarrow l}(c_{sxv}, c_{sxh}, c_{sl}) - \frac{c_{sl}}{N} p_{l \rightarrow x}(c_{sxv}, c_{sxh}, c_{sl}) \right) = 0 \quad (27)$$

where $p_{x \rightarrow l}(c_{sxv}, c_{sxh}, c_{sl})$ is the transition probability from pixel x to l for given values of c_{sxv} , c_{sxh} and c_{sl} . Using Eq. 9 the $p_{x \rightarrow l}(c_{sxv}, c_{sxh}, c_{sl})$ is specified as

$$p_{x \rightarrow l}(c_{sxv}, c_{sxh}, c_{sl}) = \min \left(1, \frac{(c_{sxv} + c_{sxh})^{c_{sxv} + c_{sxh}} (c_{sl} + 1)^{c_{sl} + 1}}{(c_{sxv} + c_{sxh} + 1)^{c_{sxv} + c_{sxh} + 1} c_{sl}^{c_{sl}}} \right). \quad (28)$$

The $c_{sxv} + c_{sxh}$ is the total number of events in central pixel x . Similarly, the $p_{l \rightarrow x}(c_{sxv}, c_{sxh}, c_{sl})$ is specified.

With the normalization condition $\pi_x^* + B\pi_l^* = 1$ the Eq. 28 was solved numerically using bisection method for π_x^* and π_l^* . The results are presented in Fig. 2 and they accurately predict the result of OE reconstruction algorithm since no approximations were used.

C. Approximations of the OE Estimate

First approximation that was used in order to show that the OE estimate approximates the ML estimate is expressed by Eq. 15. Applying the same approximation for case studied in this section and noting $c_x^* = 2N\pi_x^*$ (the expected number number of events in central pixel) and similarly $c_l^* = N\pi_l^*$ the following is obtained:

$$\frac{\pi_l^*}{\pi_x^*} = \frac{(2N\pi_x^* - 1)^{2N\pi_x^* - 1} (N\pi_l^* + 1)^{(N\pi_l^* + 1)}}{2N\pi_x^* 2N\pi_x^* N\pi_l^*} \quad (29)$$

The same as before normalization condition $\pi_x^* + B\pi_l^* = 1$ applies. Non-linear Eq. 29 was solved for π_x^* and π_l^* using bisection method and results are presented in Fig. 2. Note the the above equation do not have real number solution for $2N\pi_x^* < 1$ since v^v is undefined for $v < 0$.

Applying the second (and final) approximation expressed by Eq. 17 to the example presented here we obtain

$$\frac{\pi_l^*}{\pi_x^*} = \frac{N\pi_l^*}{2N\pi_x^*} \quad (30)$$

Using normalization condition $\pi_x^* + B\pi_l^* = 1$, the above has two solutions: ($\pi_l^* = 0, \pi_x^* = 1$) and ($\pi_l^* = 1/B, \pi_x^* = 0$). The second solution is rejected since previously it was required that $2N\pi_x^* \geq 1$. For the first solution, the value of $\pi_x^* = 1$ corresponds to activity in x equal to $2N$ which is equal to the ML estimate.

D. Results of Point Source Experiment

Figure 2 presents results for four values of B equal to 32, 64, 128, and 256. The total number of events $2N$ varied from 10 to 20,000. The figure presents point source recovery coefficient r_c . It represents the reconstructed value of emitted events in center pixel divided by the true value which was $2N$. The ML solution corresponds to $r_c = 1$.

The results show that increased accuracy in agreement of the OE and ML estimates is achieved with increased number of detected events. This is expected based on the theory described in section II-D as the approximation used in Eqs. 15 and 17 improve with higher number N . For fewer number of bins used in the experiment, the OE estimate better agrees with the the ML estimate. This again is consistent with the theory since decreasing number of pixels increases number of events per pixel leading to improved accuracy. As seen here the OE method overestimates pixels with zero activity. It follows that if the number of pixels with zero activity is smaller the accuracy should be increased. The difference between light grey solid lines which correspond to the exact OE estimate and thin solid lines are the result of the approximation expressed by Eq. 15. As expected based on the theory— the differences are significantly reduced for high number of detected counts. The results of the OE

algorithm (points with error bars in the Fig. 2) agree well with theoretical predictions derived using bi- and tri-nominal distributions (light grey lines in Fig. 2).

IV. Application to computer simulation data for SPECT

In order to directly compare the OE and ML estimates, computer simulations were used in which the reconstruction area was defined as 128×128 pixel matrix. Sixty projections over 180° were simulated each projection with 128 bins. Elements of the system matrix a_{strip} were defined as strip integrals [13]. Noiseless projection data of the phantom were generated using the system matrix a_{strip} . The data were scaled to desired total number of counts N . Two levels (low and high) of total counts were investigated corresponding to N equal to 50k, 5M counts, respectively. All of the above was done in double precision. Poisson distributed noise was added using `poidev()` function [14] in double precision. Data were reconstructed using 100,000 ML-EM iterations using system matrix a_{strip} to obtain the ML estimate. Note that the same system matrix was used to create the noiseless data and to reconstruct the image to avoid artifacts due to partial volume effects. Data were also reconstructed using the OE algorithm using 5000 iterations to reach equilibrium and another 1000 iterations to obtain OE estimates. The reconstruction region for OE was restricted by the radius slightly larger than the object. The radii were 64 (unrestricted) and 48 (restricted) pixels. For the ML only radius equal to 64 pixels was used as the ML solution is almost identical for any radius as long as it is larger than the radius of the object which was equal to 48 pixels. The OE estimates were calculated as an average of last 1000 iterations of the OE. Although the images corresponding to subsequent iterations of the OE are highly correlated this correlation does not affect the average assuming the Markov time during which samples are acquired is long [15]. The value of logarithm of the data likelihood was calculated for ML and OE for every iteration of the algorithms.

Figure 3 presents the value of the log-likelihood for low and high count experiments. Taking into account results from theory section and the section describing the point source experiment it is expected that for the high-count case the agreement with the ML should be better. It is clear that by restricting locations of the events in OE algorithm to the object only a better agreement with ML is achieved by eliminating voxels with zero activity from the reconstruction area.

Figures 4 and 5 summarize images obtained by the OE and ML methods. In Figure 4, the images filtered by the Gaussian filter with 1 pixel FWHM are presented. This was done to reduce variance of the images and to improve perception and allow better visual comparison of the images. All images were filtered with the identical filter. In general, a remarkable similarity of noise textures can be appreciated. ML-EM reconstructed image appears noisier which is confirmed by profiles in figure 5. For high count case the profiles are very similar. The OE algorithm overestimate cold regions with low number of emitted counts (Fig. 4E) compared to ML solution.

V. Application to Monte Carlo generated data for 3D list-mode PET

Monte Carlo computer simulations of positron emission tomography (PET) were used to investigate properties of the OE reconstruction compared to those of the list-mode ML-EM algorithm. A list-mode 3D PET acquisition was simulated using Simset Monte Carlo software [16] with a realistic simulated scanner geometry (44.31 radius, 15.7 axial span, 3.5 cm. BGO crystal thickness). All events were recorded in a list-mode format by registering the x, y, z coordinates of the 511 keV photon interactions with the crystal. Detector geometrical response was simulated by shifting the true interaction locations toward the central axis of the scanner to a depth in the crystal equal to the average interaction depth. Individual crystals were not simulated. The object consisted of a 30-cm diameter cylinder

filled with water (water attenuation was simulated) with 8 cold and hot spheres. The relative activities were 1, 0.2, and 5 for background, cold, and hot spheres, respectively. Sphere diameters were 4, 2, 1.5, and 1 cm. Photon attenuation was simulated assuming the cylinder was filled with water, but only the unscattered events were used in the reconstruction. Random coincidences were ignored. A total of 13.6 mln of detected coincidences was used. The reconstruction area was defined as a volume of $192 \times 192 \times 192$ matrix of voxels with size of the voxel equal to 0.3125 cm.

Two interaction locations defined a line of response (LOR). For the list-mode ML-EM reconstructions, projection operations were obtained by performing line integrals with integration step equal to half size of the voxel. Similarly, backprojection operations were performed for each LOR. For OE new event locations (Step 3 of the algorithm) were randomly generated on LORs in restricted area which was defined as cylinder with radius of 64 voxel sizes.

Attenuation correction was performed for both the list-mode ML-EM and the OE algorithms. For the ML-EM the standard algorithm was used to perform attenuation correction. For OE attenuation correction was performed as described by [7]. For the case of PET it consisted of dividing the image obtained by the OE algorithm (Fig. 4C) by the sensitivity image (Fig. 4A).

Figure 6 presents images corresponding to the ML and OE estimates which are similar with the OE image being slightly smoother than ML image. Again, the noise textures are identical. All noise correlation artifacts are present in images obtained by both approaches. Note that this is true for images that are attenuation corrected (Fig. 6(B) and (D)). Although only a single slice is presented in Fig. 6, all other slices shows the same similarities. The reconstruction using single computational thread on Intel Xeon E5520 2.27 GHz processor was equal to 54 seconds per iteration and 10 seconds per iteration for ML and OE approaches, respectively. Neither algorithm was optimized for the computational speed.

VI. Discussion

The OE ensemble approach provides a new theory and a new view on the statistical reconstruction of photon limited tomographic data. Aside from theoretical novelties, the main motivation for development of these methods are practical applications. Iterative methods based on image likelihood are currently used as methods of choice in image reconstruction of photon limited data. It is conceptually straightforward to construct an algorithm based on the data likelihood for reconstruction of data acquired by a high-complexity system. Unfortunately, for large systems with very high number of ways of how the event can be detected and/or with non-sparse system matrices ML-based methods are not practical from the point of view of computing time due to considerable time needed to complete projection and backprojection operations. Approaches to speeding the reconstruction using precomputing of the system matrix (see eg [17]) are not applicable to such system because of the enormous sizes of such matrices even when they are compressed.

These high-complexity applications include reconstruction of the Compton camera data with modeling of energy and spatial resolutions of detectors and Doppler broadening. It has been reported that even without modeling of the energy resolution, the OE achieve approximately 10 fold decrease the in reconstruction time compared to ML [18]. Adding modeling of the energy and spatial resolutions and Doppler broadening will likely result in additional orders of magnitude speed up factors of the OE over the ML. The difference in computing time comes from the fact that for the ML the projection and backprojection operations are computationally very expensive when probability density function (pdf)

describing energy and spatial resolutions and Doppler broadening are taken into account (many voxels need to be summed/updated in projection/backprojection operation for every detected event). For the OE, the pdf are used only when moving events to a new voxel (steps 3 or 4 of the algorithm) which is quite straightforward and requires only a single evaluation of the pdf (always only single voxel is evaluated/updated regardless the complexity of the pdf). In fact, using complex pdf results in only a slight increase in computing time over the approach in which a simplistic pdf (eg. half-cone surface response for the Compton camera [19], [20]) is used. It is possible that for systems modeling of entire pdf, the OE approach will require longer time to reach equilibrium as is the case for general Markov Chains approaches [21], but it is speculated that this increase will not be significant compared to gains in reconstruction speeds achieved over the iterative approaches based on data likelihood. Aspects of computational efficiency and speed will be investigated in details in future work.

Another challenging area in which the OE can be used is SPECT. Typically, based on photo multiplier tube (PMT) currents, the localizations of events are determined based on the Anger logic. Obviously, PMT currents are random variables and so is the event localization that are calculated based on these currents. Inclusion of this information in the ML algorithm is quite challenging from the computational point of view. Again, it is expected that OE will have great benefits in terms of ease of implementation and computation time especially with additional complexity of modeling of the collimator response and model-based Compton scatter.

Final example from non-exhaustive list of prospective applications, but potentially of a high significance, is PET and time-of-flight PET. In current applications Compton scatter estimation is calculated based on estimates of activity. Then, the Compton scatter estimate is effectively subtracted from data (modeled in projection but not in backprojection step of iterative algorithms). It is conceptually simple to add Compton scatter to definition of the system matrix and then use the same system matrix in projection and backprojection operations, yet it is impractical with current computational capabilities. With the OE, scattered photons can be included in the reconstruction with a relative ease. The benefit of using scattered photons in image estimation is not fully studied but can be substantial as shown theoretically using the Cramer-Rao bound calculation for SPECT [22] and modeling of high order of scatters in projection and backprojection for PET [23]. Although, the details of including scattered photons in OE approach are not provided in this work it is an active area of research in our laboratory and will be described in the future.

As shown in this work, the OE estimate is similar to the ML estimate especially for voxels that many more than 1 number of events. It is important that for increased accuracy only the volume of the object is taken into account when using OE algorithm to eliminate voxels with zero activity that are outside the object. This can be easily achieved in realistic imaging conditions since for a large majority of ET scans the attenuation maps are measured either by X-Ray CT or some other transmission sources. As shown in the section III the OE algorithm does not do well in terms of reconstruction of point sources on cold backgrounds especially for low number of detected events. Fortunately this is never an issue in practical applications since the warm background improves quantitation considerably to the point that differences between OE and ML approaches are negligible.

Obviously, question arises about areas inside the volumes which are cold with no events emitted. As it was shown OE overestimates activity in these regions compared to ML estimate although the variance appears to be smaller. Since photon limited data are obtained in molecular imaging, the tracers are distributed using blood stream and at least a small fraction events emitted per voxel are always expected. We saw in section IV that already

small activity per voxel makes the reconstructions of OE better agreeing with ML estimates. Assuming worse case scenario that part of the reconstruction volume is a cold area with with no emitted events, the relative bias of OE will be high, however, the absolute bias will likely not significant for clinical findings. Nonetheless, a caution should be used for clinical tasks involving imaging of cold (no emitted events) areas in the image.

As shown in this work the OE estimate approximates the ML estimate. However, the ML estimate is never used in practice due to large variances in the reconstructed images. For the OE to become a practical method, regularization approaches need to be developed. Currently, three variance reduction approaches for the OEs are considered. The most straightforward approach to reduce variance is to use post-filtering. As shown in this work however, the OE estimate is close to ML estimate thus the high variance (although smaller than for ML) may generate substantial errors in the OE estimate. Another approach is to utilize *a priori* information which will modify state probabilities described by Eq. 6. Although straightforward extension from the iterative reconstruction methods this approach will considerably diminish performance of the OE reconstructions as the prior information would have to be evaluated in every move of the OE algorithm.

The other regularization approach that can be used with OE which does not significantly hinder reconstruction speeds is to use a regularization parameter β [24] which modifies transition probabilities (Eq. 9) as follows:

$$\frac{p_{n(i \rightarrow i')}(c_{si}, c_{si'})}{p_{n(i' \rightarrow i)}(c_{si}, c_{si'})} = \left(\frac{\alpha_{ni'}(c_{si} - 1)^{c_{si}-1}(c_{si'} + 1)^{c_{si'}+1} / \epsilon_{i'}}{\alpha_{ni}c_{si}^{c_{si}}c_{si'}^{c_{si'}} / \epsilon_i} \right)^\beta \quad (31)$$

Such defined transition probabilities obey detailed balance and ergodicity conditions – a requirement for achieving steady state. The effect on the reconstruction image is quite simple. For $\beta < 1$, more states contribute to the average since $\beta < 1$ effectively increase chance that low probability states are selected by the OE algorithm. An average consisting of more uncorrelated states would of course have lower variance. The value of beta controls the regularization level. Note that typical bias–variance trade–off occurs. The β regularization approach briefly described here will be studied in the future. The list of these three methods is probably not exhaustive and it is likely that other ingenious methods of regularization of OE estimator exist.

Interestingly, in the OE approach the correction for Compton scatter and randoms in PET can be implemented in a different manner as done for the standard iterative reconstruction where the estimates of scatter and randoms rates are added in the projection operation. In the OE, the conditional acceptance of events can be used. Simply, in step 2 of the algorithm, the event under consideration is either accepted or rejected based on estimated rates of trues, scatter, and randoms for the given LOR. If accepted, the event is processed the same way as described in this paper. If rejected, it is simply temporarily removed from the calculation. Next time rejected event is considered by the OE algorithm the status of the event can be changed to being accepted.

VII. Summary

The properties of the OE estimate were investigated. It was shown that OE estimate is equal to the expectation of maximized activities for complete–data space using Poisson likelihoods. Next, it was demonstrated that approximated–OE estimate achieves maximum data likelihood. The accuracy of these approximations increases with the number of events emitted per voxel therefore it increases with the number of detected counts N . It was found

that in order to achieve a better agreement of OE and ML estimates, the OE reconstructions should be performed restricting locations of origins to the object (eliminating voxels with zero activity) which can be achieved in real imaging situations if attenuation maps are acquired. This can also be done by first reconstructing the object and then using thresholding to obtain the volume of the object and using it in the OE reconstruction as the definition of the object volume. The effects of the approximations were studied in section III using a simple tomographic system for which the exact OE and ML estimates were calculated. In sections IV and V, the OE and ML methods were compared for SPECT with no attenuation modeled and for Monte Carlo generated data for 3D PET with detector response and attenuation. For both approaches for SPECT and attenuation corrected 3D PET images obtained by OE and ML were very similar. The similarity was assessed by examining the noise textures obtained by both method in sections IV and V. The OE approaches tends to slightly overestimate cold areas in the image. It was speculated that for practical applications this bias is negligible, but requires further studies.

Acknowledgments

The author would like to thank Drs. Stephen C. Moore, Andrei Andreyev, Anna Celler, and Marie Foley Kijewski for helpful discussions and help on the manuscript. Author also would like to acknowledge Dr. Dan Kadmas for useful insights concerning section II-C.

The work was supported by American Heart Association, Scientist Development Grant 0735328N, and National Institutes of Health grants R21 CA123057, R01 EB001989, and R01 EB 000802

References

1. Qi J, Leahy RM. Iterative reconstruction techniques in emission computed tomography. *Phys Med Biol.* 2006; 51(15):R541–78. [PubMed: 16861768]
2. Defrise M, Gullberg GT. Image reconstruction. *Phys Med Biol.* 2006; 51(13):R139–54. [PubMed: 16790900]
3. Gabor, TH. *Advances in Pattern Recognition. 2.* Hoboken: Springer; 2002. *Fundamentals of Computerized Tomography: Image Reconstruction from Projections.* ser
4. Shepp L, Vardi Y. Maximum likelihood reconstruction for emission tomography. *Medical Imaging, IEEE Transactions.* 1982; 1:113–122.
5. Lange K, Carson R. “em” reconstruction algorithms for emission and transmission tomography. *J Comput Assist Tomogr.* 1984; 8(2):306–16. [PubMed: 6608535]
6. Carson RE. A maximum likelihood method for region-of-interest evaluation in emission tomography. *J Comput Assist Tomogr.* 1986; 10(4):654–63. [PubMed: 3488338]
7. Sitek A. Representation of photon limited data in emission tomography using origin ensembles. *Phys Med Biol.* 2008; 53(12):3201–3216. [PubMed: 18506073]
8. Sitek A, Huesman RH, Gullberg GT. Tomographic reconstruction using an adaptive tetrahedral mesh defined by a point cloud. *Medical Imaging, IEEE Transactions.* 2006; 25(9):1172.
9. Matej S, Herman GT, Narayan TK, Furuie SS, Lewitt RM, Kinahan PE. Evaluation of task-oriented performance of several fully 3d pet reconstruction algorithms. *Phys Med Biol.* 1994; 39(3):355–67. [PubMed: 15551585]
10. Metropolis N, Rosenbluth A, Rosenbluth M, Teller A, Teller E. Equation of state calculations by fast computing machines. *J Chem Phys.* 1953; 21:1087–1092.
11. Hastings W. Monte carlo sampling methods using markov chains and their applications. *Biometrika.* 1970:97–109.
12. Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the em algorithm. *Journal of Royal Statistical Society.* 1977; 39(1):1–38.
13. Huesman, RH.; Gullberg, GT.; Greenberg, WL.; Budinger, TF. Tech Rep Pub. E.O. Lawrence Berkeley National Laboratory; Reclbl library. donner algorithms for reconstruction tomography; p. 214

14. Press, W.; Teukolsky, S.; Vetterling, W.; Flannery, BP. Numerical Recipes in C++. New York: Cambridge University Press; 2002.
15. Newman, M.; Barkema, G. Monte Carlo Methods in Statistical Physics. New York: Oxford University Press; 1999.
16. Harrison, R.; Vannoy, S.; Haynor, D.; Gillispie, S.; Kaplan, M.; Lewellen, T. Preliminary experience with the photon history generator module of a public-domain simulation system for emission tomography. Nuclear Science Symposium and Medical Imaging Conference, 1993., 1993 IEEE Conference Record; 1993. p. 1154-1158.
17. Scheins JJ, Herzog H. Pet reconstruction software toolkit -presto a novel, universal c++ library for fast, iterative, fully 3d pet image reconstruction using highly compressed, memory-resident system matrices. Oct.2008 :4147–4150.
18. Andreyev A, Sitek A, Celler A. Stochastic image reconstruction method for compton camera. Med Phys. Oct.2010 p. accepted.
19. Singh M. An electronically collimated gamma camera for single photon emission computed tomography. part i: Theoretical considerations and design criteria. Med Phys. 1983; 10(4):421–7. [PubMed: 6604216]
20. Singh M, Doria D. An electronically collimated gamma camera for single photon emission computed tomography. part ii: Image reconstruction and preliminary experimental measurements. Med Phys. 1983; 10(4):428–35. [PubMed: 6604217]
21. Desai MP, Rao VB. On the convergence of reversible markov chains. SIAM J Matrix Anal Appl. 1993; 14(4):950–966.
22. Sitek, A.; Foley Kijewski, M. All-count reconstruction: Effects of energy band width. J Nucl Med; MEETING ABSTRACTS; May 2009; 2009. p. 1503
23. Tamal M, Reader A, Markiewicz P, Julyan P, Hastings D. Noise properties of four strategies for incorporation of scatter and attenuation information in pet reconstruction using the em-ml algorithm. Nuclear Science, IEEE Transactions on. Oct; 2006 53(5):2778–2786.
24. Sitek, A. Image reconstruction in emission tomography using canonical origin ensembles. Nuclear Science Symposium Conference Record (NSS/MIC), 2010 IEEE; 2010. (accepted)

Appendix A. Approximation of detailed balance equation

A subset of states $\{S\}$ of size I^{N-1} is identified for which the origin of event n is located in voxel i . There is a corresponding disjoint subset of states $\{S'\}$ for which event n is located in i' . The one-to-one correspondence between these two subsets exists where two corresponding states $s \in \{S\}$ and $s' \in \{S'\}$ differ only by the location of the origin of the event n . The detailed balance equation (enforced equilibrium condition) between these two states is

$$\Pi(s)p_{s \rightarrow s'} = \Pi(s')p_{s' \rightarrow s}, \quad (32)$$

where $p_{s \rightarrow s'}$ is the transition probability from s to s' . Summing for all states in the subsets $\{S\}$ and $\{S'\}$

$$\sum_{s \in \{S\}} \Pi(s)p_{s \rightarrow s'} = \sum_{s' \in \{S'\}} \Pi(s')p_{s' \rightarrow s} \quad (33)$$

Now the crucial approximation is made. Assuming that a large majority of states with the highest value of $\Pi(s)$ have similar $p_{s \rightarrow s'}$ the above is approximated by:

$$\frac{\sum_{s \in \{S\}} P_{s \rightarrow s'}}{I^{N-1}} \sum_{s \in \{S\}} \Pi(s) \approx \frac{\sum_{s' \in \{S'\}} P_{s' \rightarrow s}}{I^{N-1}} \sum_{s' \in \{S'\}} \Pi(s') \quad (34)$$

From the definition $\sum_{s \in \{S\}} \Pi(s) = \pi_{ni}^*$ and $\sum_{s' \in \{S'\}} \Pi(s') = \pi_{ni'}^*$. Noting that $\frac{\sum_{s \in \{S\}} P_{s \rightarrow s'}}{I^{N-1}}$ and $\frac{\sum_{s' \in \{S'\}} P_{s' \rightarrow s}}{I^{N-1}}$ are the average transition rates of moving origin of event n from i to i' (from definition of the subsets) the first approximation in Eq. 15 is derived.

Appendix B. Derivation of condition for ML solution

It is assumed that number of detected events in projection bins is independent Poisson distributed random variable. For this case the likelihood function L is of the form:

$$L = \prod_{k=1}^K \frac{\widehat{g}_k^{g_k} e^{-\widehat{g}_k}}{g_k!}. \quad (35)$$

Taking the logarithm of likelihood function and omitting constant terms:

$$\log L = \sum_{k=1}^K g_k \log \widehat{g}_k - \sum_{k=1}^K \widehat{g}_k. \quad (36)$$

The first summation of the above equation can be rewritten for the list-mode data which leads to:

$$\log L = \sum_{n=1}^N \log \widehat{g}_n - \sum_{k=1}^K \widehat{g}_k. \quad (37)$$

Function described by Eq. 37 is globally convex and has a single global maximum. Thus, the first derivative vanishes at the extremum. Taking this into consideration and the fact the $\widehat{g}_n = \sum_{i=0}^I \alpha_{ni} \widehat{f}_i^{ML}$ and sensitivity matrix defined by Eq. 1 $\varepsilon_i = \sum_{k=1}^K \alpha_{ki}$ the sufficient condition for the list-mode maximum likelihood is obtained:

$$0 = \frac{d \log L}{d \widehat{f}_i^{ML}} = \sum_{n=1}^N \frac{\alpha_{ni}}{\sum_{i'=1}^I \alpha_{ni'} \widehat{f}_{i'}^{ML}} - \varepsilon_i. \quad (38)$$

Appendix C. Equations for Multinomial Distributions

In section III formulas for the bi-nominal and the tri-nominal distributions were used. Binomial distribution $P_{\text{Bin}}(a, Z, p_a)$ describes probability of number of successes a in Z trial ($a \leq Z$) experiment where a chance for successes in single trial is p_a . The $P_{\text{Bin}}(a, Z, p_a)$ is described by:

$$P_{\text{Bin}}(a, Z, p_a) = \frac{Z!}{a!(Z-a)!} p_a^a (1-p_a)^{Z-a} \quad (39)$$

Tri-nominal distribution $P_{\text{Trin}}(a, b, Z, p_a, p_b)$ describes a probability distribution where for the Z trial experiment the outcome can be either of three possibilities each with a probability p_a , p_b , and $(1 - p_a - p_b)$.

$$P_{\text{Trin}}(a, b, Z, p_a, p_b) = \frac{Z!}{a!b!(Z-a-b)!} p_a^a p_b^b (1-p_a-p_b)^{Z-a-b} \quad (40)$$

Fig. 1.

The setup of the point source thought experiment. Two projections were acquired at the vertical and horizontal position. Reconstruction area is of $(B+1) \times (B+1)$ size. For the example shown in this figure $B=4$. Each projection has $B+1$ projection bins. Only central bin detects N events and values detected at all other bins are 0. While in equilibrium for the light shaded pixels the probability that an event is located in given bin is π_l^* and is the same for all light shaded bins and all detected photons. For dark pixel this probability is π_x^* and is the same for all $2N$ detected photons.

Fig. 2.

The Figure presents values of the recovery coefficient for a point source as a function of N for the exact prediction of OE estimate (solid light grey curves). Going from left, four curves correspond to B equal to 32, 64, 128, and 256. Points with error bars represent results of the OE algorithm and error bars represent standard deviation of the value of r_c while in equilibrium. Prediction of the r_c for case of approximation of the OE using Eq. 15 is indicated by solid lines. Line for value of $r_c = 1$ corresponds to the ML estimate.

Fig. 3.

The log-likelihood as a function of iterations for the ML and OE approaches. Top graph corresponds to 50,000 events (low count experiment) and a bottom graph to 5,000,000 events (high count experiment).

Fig. 4.

Results of the reconstructions using the ML and OE approaches. The left column (A, C and E) corresponds to results for $N=50k$ and right column (B, D, and F) corresponds to $N=5M$. To achieve better perception and appreciate noise textures the original images were filtered using Gaussian filter with 1 pixels of FWHM. The greyscale was also adjusted so the background is displayed with good contrast. The same color scale was applied to images reconstructed from 50k (left column) and for 5M (right column) counts. The OE reconstructed images in second row correspond to entire reconstruction area (64 pixel radius), and bottom row corresponds to area restricted to the object (48 pixels).

Fig. 5.

Profiles for images presented in Fig. 4. Left column corresponds to $N = 50k$ and right column correspond to $N = 5M$. Upper 2 rows represents vertical profiles positioned at the center of the image, and bottom row corresponds to horizontal profiles through the center of the image. Note, that different scale for used for vertical profiles to visualize high and low count emitting regions of the image. True, ML, OE, and restricted OE profiles are presented.

Fig. 6.

A center slice through the $192 \times 192 \times 192$ reconstruction volume for list-mode PET. A vertical line in (A) corresponds to a profile through the center of the shown slice. (A) shows the sensitivity image that was used for the ML-EM and OE approaches. Low sensitivity in the center is due to photon attenuation (B) shows the result of 10,000 iterations of ML-EM with attenuation correction. (C) shows of the result of the OE algorithm. Since the image represents estimated number of events that were emitted and detected per voxel, lower number of events in the center is seen. (D) shows image (C) divided by (A) which corresponds result of the OE (for this case with attenuation correction).