

Proceedings

# Bayesian Approach with Entropy Prior for Open Systems <sup>†</sup>

Natalya Denisova

Institute of Theoretical and Applied Mechanics, 630090 Novosibirsk, Russia; NVDenisova2011@mail.ru

<sup>†</sup> Presented at the 39th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, Garching, Germany, 30 June–5 July 2019.

Published: 12 November 2019

**Abstract:** The Bayesian approach Maximum a Posteriori (MAP) is discussed in the context of solving the image reconstruction problem in nuclear medicine: positron emission tomography (PET) and single photon emission computer tomography (SPECT). Two standard probabilistic forms, Gibbs and entropy prior probabilities, are analyzed. It is shown that both the entropy-based and Gibbs priors in their standard formulations result in global regularization when a single parameter controls the solution. Global regularization leads to over-smoothed images and loss of fine structures. Over-smoothing is undesirable, especially in oncology in diagnosis of cancer lesions of small size and low activity. To overcome the over-smoothing problem and to improve resolution of images, the new approach based on local statistical regularization is developed.

**Keywords:** image reconstruction problem; Bayesian Maximum a Posteriori approach; entropy prior probability; global regularization; local statistical regularization; open systems

---

## 1. Introduction

Emission tomography techniques, including (PET) and single photon emission computer tomography (SPECT), produce images which are used for clinical diagnosis. Image quality depends substantially on applied reconstruction methods. The deterministic filtered back projection (FBP) method was used for image reconstruction up to the 90-th. Later, statistical iterative methods based on the maximum likelihood principle, entered the nuclear medicine. Statistical algorithms, such as the Maximum Likelihood Expectation Maximization (MLEM) and its accelerated version - Ordered Subsets Expectation Maximization (OSEM)—take into account the stochastic properties of the observed data and provide more accurate images in comparison to FBP method. At the moment, OSEM is the standard algorithm in PET and SPECT systems all over the world. However, OSEM algorithm has fundamental limitations in solving reconstruction problems and unexpected artifacts may arise in the obtained images. These limitations are due to the fact that from the mathematical point of view, image reconstruction belongs to the class of ill-posed inverse problems and regularization is necessary for its correct solution. Regularization is the process of introducing an additional a priori assumption about the solution to obtain a well behaved inverse [1]. Initially, the regularization method was developed by using the deterministic form of prior information. However, the deterministic form of prior information has limitations with the drawback to lead to excessively over-smoothed solutions. The stochastic nature of the observed data requires the probabilistic approach for assigning prior information. V. Turchin suggested to use the Bayesian method of Maximum a Posteriori (MAP) for solving ill-posed problems with stochastic data, naming this approach ‘statistical regularization’ [2]. The statistical regularization method introduces the prior information in the form of a probability distribution. Two basic probabilistic forms of prior information, widely used in reconstruction tomography, are: Gibbs prior and entropy prior. Both deterministic and statistical regularization methods were developed as ‘global regularization’, in

which a single parameter controls the solution in the whole solution's area. It was expected, that the regularized MAP algorithms should provide more accurate reconstruction of fine structures in comparison to the non-regularized OSEM. However, in [3], rather minor differences between the images obtained by the OSEM algorithm and MAP algorithm with Gibbs prior (MAP-Gibbs) were found. In [4], numerical simulations have shown that by using the entropy-based MAP algorithm (MAP-ENT), the reconstruction errors decrease monotonically with excellent stability, in contrary, the OSEM algorithm behavior was unstable. The OSEM reconstructed image was obtained by stopping the iteration process at the point of minimal error and similarly to the resulting MAP-ENT image. We assume that the cause of the minor difference between OSEM and MAP images is due to the global regularization method, applied in both MAP algorithms. Global regularization smoothes the solution too strong and therefore fine structures may be over-smoothed or lost. Local regularization is needed to improve image quality. The idea of local regularization for statistical MAP approach has not been considered in literature.

The aim of this paper is the theoretical analysis of local regularization method in the frame of statistical Bayesian MAP approach.

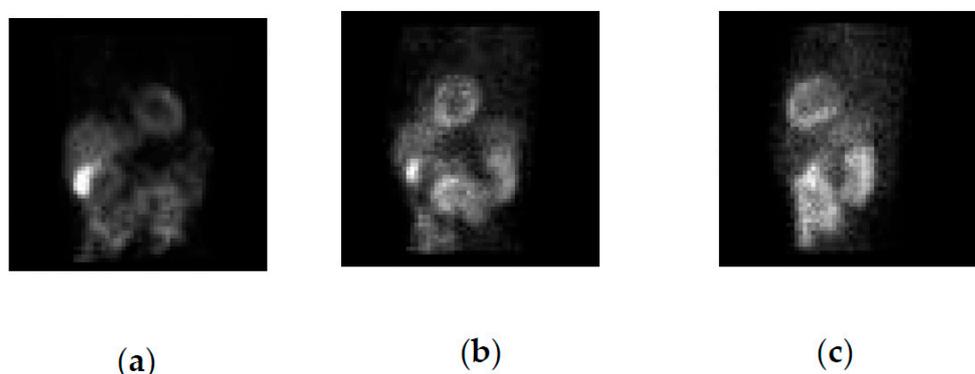
## 2. Methods

### 2.1. Bayesian Approach for Solving Tomographic Problems in Nuclear Medicine

In SPECT and PET diagnostic procedures, a patient is administered a radiolabeled pharmaceutical which is distributed with the concentration  $n(q)$  in various regions of the body ( $q$  is the space coordinate). The function  $f(q)$  describes the density of gamma photons produced through radioactive decay. It is assumed that  $f(q)$  is a random value which follows the Poisson distribution with mean  $\bar{f}$  which is proportional to the radiopharmaceutical concentration  $n$ . The image reconstruction problem is presented as linear equation:

$$Af = g \tag{1}$$

$A$  is a system matrix which describes data acquisition process and  $g$  are registered stochastic data. Gamma photons are emitted by the radiopharmaceutical and are registered by gamma camera. Gamma camera rotates around the patient body and collects gamma photons from different angles. The registered raw data are called projection data. In myocardial perfusion SPECT imaging the data are obtained from 32 or 64 views. An example of clinical raw projection data is shown in Figure 1. The data of 3 from total 64 views are shown.



**Figure 1.** Clinical SPECT projection data. The data are demonstrated for three selected views: anterior (a), left anterior oblique (b) and left lateral (c) projections.

The data are obtained in Meshalkin National Medical Research Center (Novosibirsk) by using GE Infinia Hawkeye SPECT/CT system.

The data are assumed to be distributed according to a Poisson law:

$$P(g | \bar{g}) = e^{-\bar{g}} \frac{\bar{g}^g}{g!} \tag{2}$$

with mean  $\bar{g}$  :

$$\bar{g} = A\bar{f} \tag{3}$$

The reconstruction problem (1) is formulated as a classical problem of mathematical statistics: what is the probability density  $P(\bar{f} | g)$  for the solution  $\bar{f}$  with given data  $g$  ? By using the Bayesian Maximum *a Posteriori* (MAP) theorem,  $P(\bar{f} | g)$  is determined as:

$$P(\bar{f} | g) = \frac{P(\bar{f})P(g | A\bar{f})}{\int P(\bar{f})P(g | A\bar{f})d\bar{f}} \tag{4}$$

where  $P(\bar{f})$  is a prior probability density function and  $P(g | A\bar{f})$  is the likelihood distribution of the observed data. The most probable MAP estimation is obtained by maximizing  $P(\bar{f} | g)$  in logarithmic form:

$$\tilde{f} = \max_{\bar{f}} \{ \ln P(\bar{f}) + \ln P(g | A\bar{f}) \} \tag{5}$$

For applications in nuclear medicine, the likelihood distribution in logarithmic form is defined in accordance to (2):

$$\ln P(g | A\bar{f}) = g \ln A\bar{f} - A\bar{f} - \ln g! \tag{6}$$

### 2.2. Maximum-Likelihood-Based Image Reconstruction Method

When a researcher does not have any prior information the simple way is to assume that all possibilities are equally probable and Bayesian estimation reduces to the Maximum Likelihood (ML) solution:

$$\tilde{f} = \arg_{\bar{f}} \max \{ \ln P(g | A\bar{f}) \} \tag{7}$$

In practical PET and SPECT applications reconstruction problem is usually discretized. The reconstruction area is assumed to be divided into  $J$  voxels in which the unknown source function  $\bar{f}$  is distributed with a discrete density  $\bar{f}_j$ . The projection data are presented in the discrete form as a system of the linear algebraic equations:

$$g_i = \sum_j A_{ij} f_j \tag{8}$$

where  $g_i$  are the measured data in the  $i$ -th detector cell, matrix  $A_{ij}$  is the system matrix describing the possibility that photon emitted in the  $j$ -th voxel will be detected in  $i$ -th 'detector'. The ML method (7) provides the following solution:

$$\tilde{f}_j^{n+1} = \frac{\tilde{f}_j^n}{\sum_i A_{ij}} \sum_i \frac{g_i A_{ij}}{\sum_j A_{ij} \tilde{f}_j^n} \tag{4}$$

$\tilde{f}_j^n$  is an estimation of  $\bar{f}_j$  on the  $n$ -th iteration step. In clinical systems, the OSEM algorithm is used. OSEM is a faster version of ML algorithm. OSEM is not a regularized algorithm, therefore, its behavior in iteration process is unstable and the resulted images are dominated by noisy artifacts. In the first iterations, the error of reconstruction decreases, then it reaches a minimum and begins to increase. At this point a reconstructed image begins to deteriorate. One may think that it is possible to obtain visually good OSEM images by stopping the iteration process at the iteration with minimal error. Nonetheless, image deterioration progresses more rapidly in the zones with lower count statistics. At stopping, a part of the image with low count statistics can become already noisy while another part of the image with high statistics did not reach its optimal resolution yet. This problem is especially important in oncology: a noise in the low count zones imitates virtual ‘hot spots’ or masks the true ‘hot spots’. Currently, stopping of OSEM algorithms is applied in standard SPECT myocardial perfusion imaging protocols. In some cases, the stopping rule may be the cause of artifacts on images.

### 2.3. The Bayesian Image Reconstruction Method with Gibbs Prior

In the literature devoted to medical emission tomography, Bayesian Maximum *a Posteriori* approach with Gibbs prior is most widely studied. This approach was justified theoretically in [5] and was first applied for SPECT by Geman and McClare in 1984 in [6]. Gibbs prior assumes the Markov Random Field (MRF) distribution for emitted photons. According to the Hammersly-Clifford theorem, the MRF has the distribution which is described by the Gibbs probability:

$$P(\bar{f}) = \frac{1}{Z} \exp(-\beta U(\bar{f})) \tag{5}$$

where  $\beta$  is an unknown parameter,  $U(\bar{f})$  is the energy function. Energy function  $U(\bar{f})$  is defined as a sum of potentials:

$$U(\bar{f}_j) = \sum_{j \in c} V_c(|\bar{f}_j - \bar{f}_i|) \tag{11}$$

where  $c$  denotes the set of all cliques type,  $V_c$  is a potential functions defined on pairwise cliques of neighboring voxels. The Bayesian MAP solution (MAP-Gibbs) of the reconstruction problem (1) is presented in this case as:

$$\tilde{f} = \arg_{\bar{f}} \max \{-\beta \sum_i \sum_{j \in c} V_c(|\bar{f}_j - \bar{f}_i|) + g \ln A\bar{f} - A\bar{f} - \ln g!\} \tag{12}$$

Gibbs prior provides the reconstruction algorithm that specifies local spatial correlations in the source to be reconstructed. A wide range of potential functions  $V_c$  have been proposed in the literature. For example, edge-preserving Geman’s approximation of the function  $V_c$  is defined as follows:

$$V_c = \frac{(\bar{f}_j - \bar{f}_i)^2}{2\delta^2 + (\bar{f}_j - \bar{f}_i)^2} \tag{13}$$

$\delta$  is edge-preserving parameter. Unfortunately, the physical nature of  $V_c$  is not very clear. Huge efforts are aimed at the development of this approach for PET and SPECT applications. However, there are unsolved problems that inhibit the implementation of these algorithms on commercial systems.

#### 2.4. Bayesian Image Reconstruction Method with Entropy Prior

Another form of prior, based on the entropy principle, was suggested by Jaynes [7–9]. Maximum-entropy (ME) approach has been successfully applied in the fields of X-ray-, radio- and gamma-astronomy, plasma tomography [10–14], but less in medicine tomography. Applied to radio-astronomical data, the maximum entropy algorithm reveals details not seen by conventional analysis, but which are known to exist. Applied to solve the problem of astronomical image restoring, the Maximum Entropy approach has demonstrated that resolved and unresolved sources can be restored with reliability. These results are of interest in the context of the similarity between astronomical images and ‘hot spots’ tumor images in nuclear oncology. In the absence of correlations the entropy-based method is superior to the Gibbs’ approach in ‘hot spots’ identifying.

The ME approach goes back to the famous works by Boltzmann and can be formulated as follows: distributions of higher entropy have higher multiplicity and can be realized in more ways. In order to apply the entropy principle to the SPECT and PET imaging problems, the reconstruction area is properly discretized into  $J$  boxes. An image is considered as a set of boxes in which a large number of emitting particles are placed. According to the combinatorial theorem, the number of different ways of filling the boxes is given by

$$W = \frac{N_0!}{\prod_j N_j!} \tag{14}$$

$j$  is a box index,  $N_j$  is the number of particles in the  $j$ -th box and  $N_0$  is the total number of particles. One usually uses the Stirling approximation to find the entropy equation in the logarithmic form. The prior probability  $P(\bar{f})$  of the unknown function is taken to be proportional to its multiplicity. The image (the number of emitted photons)  $\bar{f}_j$  is assumed to be proportional to the density of the emitting particles and, therefore, one obtains

$$\ln P(\bar{f}) = -\beta \sum_{j=1}^J \bar{f}_j \ln \bar{f}_j \tag{15}$$

where  $\beta$  is a constant associated with the relation between  $N_j$  and  $\bar{f}_j$ . Maximum *a Posteriori* estimation with entropy prior (MAP-ENT) is written in the discrete form as:

$$\tilde{f} = \arg_{\bar{f}} \max \left\{ -\beta \sum_j \bar{f}_j \ln \bar{f}_j + \sum_i [g_i \ln \sum_j A_{ij} \bar{f}_j - \sum_j A_{ij} \bar{f}_j - \ln g_i!] \right\} \tag{16}$$

Unlike Gibbs prior, Entropy form does not introduce correlations in the images, beyond those which are required by the data. It is necessary to make the following remark concerning the maximum entropy approach. There are two basic approaches using the entropy prior: the Maximum *a Posteriori* algorithm with entropy prior and the constrained Maximum Entropy algorithm. A difference between these approaches is as follows: in the MAP algorithm, the data enters as a part of quantity to be maximized,  $P(\bar{f} | g)$ . With the Maximum Entropy method, the data are added via Lagrange multipliers exterior to the entropy prior  $P(\bar{f})$  to be maximized. The Maximum Entropy yields the smoothest solution among all solutions satisfying to given data. A degree of smoothness of the MAP solution is regulated by the regularization parameter.

#### 2.5. Bayesian Image Reconstruction Method Based on Open System Theory

Gibbs and entropy forms of prior information result in global regularization method with a single regularization parameter. Global regularization produces too smoothed solution therefore fine structures may be lost on images. Local regularization is needed to improve image quality. The idea of local regularization for statistical MAP approach is considered in this Section. This idea is

based on the open systems theory. Open systems are the systems which can be exchanged with environment by energy, matter and information and, from this point of view, emitting objects are open systems. In SPECT and PET diagnostic procedures a radiopharmaceutical is injected into a patient body. Different organs have different metabolism rate for the injected radiopharmaceutical. In the same organ, healthy and ill tissues can also have different metabolism rate. Different radiopharmaceutical accumulation in the organs is associated with a different metabolic rate, which is considered here as a control parameter 'a'. We consider the accumulative model that describes a final steady-state non-equilibrium spatial distribution of radiopharmaceutical in a patient body. Due to radioactive decay, radiopharmaceutical emits gamma photons, so a patient body can be considered as an emitting open system. We assume that processes of nuclear excitation and de-excitation occur at the same rate during all the time of patient examination procedure. So, the distribution of radiopharmaceutical particles in a patient body can be considered as a steady-state open system. The steady-state of the system is defined by a corresponding control parameter. Due to low radiopharmaceutical concentration we can consider an open system of  $N$  non-interacting classical particles (ideal gas) in the volume  $V$  with the energy  $E$ . In an equilibrium system, particles are uniformly distributed throughout the volume. Non-uniform spatial distribution of particles can occur in the non-equilibrium steady-state systems. Macroscopic states of an ideal gas can be described through the possible energy distribution of an individual particle. Following [15], in classical case the entropy can be written as:

$$S = - \int g n(q, p, a) \ln n(q, p, a) \frac{dpdq}{\Delta} + \text{const} \tag{17}$$

$q, p$  are spatial and pulse coordinates,  $\Delta = (2\pi\hbar)^6$ ,  $g$  is a weight factor associated with different nuclear states of particles. Assuming that energy (pulse) distributions are the same at each point in a discrete physical space, we obtain

$$S = - \sum_j g n(a)_j \ln n(a)_j + \text{const} \tag{18}$$

$n_j(a)$  is radiopharmaceutical concentration in the  $j$ -th voxel of physical space.

For simplicity, let us focus on one organ, for example, on a liver with the area of healthy tissue and some area of ill tissue. Suppose only two cases with the concentration of radiopharmaceutical particles in the liver  $n(a_0)$  and  $n(a_1)$ . Calculation of entropy can be performed separately for each of these two areas in accordance with the Equation (18). However, definition of total entropy by using the usual Boltzmann formula is not possible because the mean energies of these two states are different. According to Klimontovich S-theorem [16], we can define the total entropy after renormalizing of energies:

$$\langle H(n(a_0)) \rangle = \langle H(n(a_1)) \rangle \tag{19}$$

$H$  is an Hamiltonian function. Integrating (17) by pulses  $p$  under condition (19) one obtains the renormalized entropy expressions for the healthy and ill liver areas. Taking into account that mean  $\bar{f}$  is proportional to the radiopharmaceutical concentration, we can write the total entropy as:

$$S = - \sum_k \beta_k \sum_{j \in C_k} \bar{f}_j(a_k) \ln \bar{f}_j(a_k) \tag{20}$$

the index  $k$  defines the subsystems with different parameter  $a$  (different radiopharmaceutical concentration). Maximum *a Posteriori* estimation with entropy prior based on open system theory (MAP-OPEN-ENT) is written in the discrete form as:

$$\bar{f} = \arg_{\bar{f}} \max \{ - \sum_k \beta_k \sum_{j \in C_k} \bar{f}_j(a_k) \ln \bar{f}_j(a_k) + \sum_i [g_i \ln \sum_j A_{ij} \bar{f}_j - \sum_j A_{ij} \bar{f}_j - \ln g_j !] \} \tag{21}$$

In solution of SPECT and PET reconstruction problems by using Bayesian approach, the expression (21) leads to local regularization with local regularization parameters  $\beta_k$ .

### 3. Conclusions

The standard approach for solving the ill-posed inverse problem in image reconstruction is used in the form of ‘global regularization’, in which a single parameter controls the solution. However, in SPECT and PET imaging, global regularization leads to over-smoothed images and loss of fine structures. Over-smoothing is undesirable, especially in oncology in diagnosis of cancer tumors of small size and low activity. To overcome the over-smoothing problem and to improve resolution of images, the new approach based on local statistical regularization is developed in this work. The theoretical justification of the new method in which the value of the regularization parameter is defined on the base of open system entropy is performed.

**Acknowledgments:** This work was partly supported by Russian Foundation for Basic Research (grant No. 17-52-14004).

**Conflicts of Interest:** The author has no competing interests to disclose.

### References

1. Tikhonov, A.N. Solution of incorrectly formulated problems and the regularization method. *Sov. Math. Doklady* **1963**, *4*, 1035–1038.
2. Turchin, V.F. A solution of Fredholm equation in statistical ensemble of smooth functions. *J. Comput. Math. Math. Phys.* **1967**, *6*, 1270. (In Russian)
3. Nuyts, J.; Fessler, J.A. A penalized-likelihood image reconstruction method for emission tomography, compared to post-smoothed maximum-likelihood with matched spatial resolution. *IEEE Trans. Med. Imaging* **2003**, *22*, 1042–1052.
4. Denisova, N.V.; Terekhov, I.N. A study of myocardial perfusion SPECT imaging with reduced radiation dose using maximum likelihood and entropy-based maximum a posteriori approaches. *Biomed. Phys. Eng. Express* **2016**, *2*, 055015.
5. Geman, S.; Geman, D. Stochastic Relaxation, Gibbs Distributions, and the Bayesian restoration of Images. *IEEE Trans. Pattern Anal. Mach. Intell.* **1984**, *6*, 721–741.
6. Geman, S.; McClure, D. Bayesian image analysis: An application to single photon emission. *Proc. Stat. Comput. Sect.* **1985**, 12–18. Available online: <https://discover.libraryhub.jisc.ac.uk/search?q=subject%3A%20Sampling%20%5BStatistics%5D%20Congresses.&rn=11&for=ncl> (accessed on 12 November 2019).
7. Jaynes, E.T. Prior Probabilities. *IEEE Trans. Syst. Sci. Cybern.* **1968**, *4*, 227–241.
8. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630.
9. Jaynes, E.T. Information theory and statistical mechanics II. *Phys. Rev.* **1958**, *108*, 171–190.
10. Gull, S.F.; Daniell, G.J. Image reconstruction from incomplete and noisy data. *Nature* **1978**, *272*, 686.
11. Skilling, J. The axioms of maximum entropy. In *Maximum Entropy and Bayesian Methods in Science and Engineering*; Erickson, G.J., Smith, C.R., Eds.; Springer: Heidelberg, Germany, 1988; Volume 1, pp. 173–188.
12. Frieden, B.R. Restoring with maximum likelihood and maximum entropy. *J. Opt. Soc. Am.* **1972**, *62*, 511–518.
13. Von der Linden, W. Maximum Entropy data analysis. *Appl. Phys. A* **1995**, *60*, 155.
14. Denisova, N. Plasma diagnostics using computed tomography method. *IEEE Trans. Plasma Sci.* **2009**, *37*, 502.
15. Landau, L.D.; Lifshitz, E.M. *Statistical Physics Theoretical Physic*. Pergamon Press: Oxford, UK, 1964. Volume 5.
16. Klimontovich Y.L. S-theorem. *Z. Phys.* **1987**, *66*, 125–127.

