LLNL-JRNL-523993

LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

Reconstruction Algorithm for Point Source Neutron Imaging through Finite Thickness Scintillator

H. Wang, V. Tang, J. F. McCarrick, S. Moran

January 24, 2012

Nuclear Instruments and Methods in Physics Research
Section A

Disclaimer

Reconstruction Algorithm for Point Source Neutron Imaging through Finite Thickness Scintillator

Han Wang, Vincent Tang, James McCarrick, and Shane Moran
Lawrence Livermore National Laboratory

1 Abstract

A new inversion algorithm based on the maximum entropy method (MEM) is proposed to remove unwanted effects in fast neutron imaging which result from an uncollimated source interacting with a finitely thick scintillator. The algorithm takes as an input the image from the thick scintillator (TS) and the radiography setup geometry. The algorithm then outputs a restored image which appears as if taken with an infinitesimally thin scintillator (ITS). The inversion is accomplished by numerically generating a probabilistic model relating the ITS image to the TS image and then inverting this model on the TS image through MEM. Algorithm details as well as numerical results using MCNP simulated images are presented. This reconstruction technique can reduce the exposure time or the required source intensity without undesirable object blurring on the image by allowing the use of both thicker scintillators with higher efficiencies and closer source-to-detector distances to maximize incident radiation flux. The technique should also be applicable to high energy gamma or x-ray radiography using thick scintillators.

Keywords: Maximum Entropy Method; Support Vector Machines; Fast Neutron Radiography; MCNP; Cone Beam Effect

2 Introduction

Fast neutron imaging is an active area of research as it offers unique imaging modalities compared with traditional x-ray and thermal neutron imaging, such as the ability to nondestructively discern features in low-Z objects shielded by thick high-Z materials [1]. In digital fast neutron imaging for example, fast neutrons are passed through a target onto a scintillator whose light is collected by a CCD camera. Scintillator thicknesses of multiple centimeters are required to detect MeV level neutrons with viable efficiencies. Collimated neutron beams are thus typically used because un-collimated and divergent beams will induce a cone beam effect in the resulting image due to both the finite thickness of the target and the scintillator itself. However, collimating a neutron source through increased source distance-to-target drastically reduces the neutron flux incident on the target and results in significantly longer imaging times. Here, we present one solution for uncollimated neutron imaging which removes the cone beam effect caused by the finite thickness of the scintillator via post-processing imaging reconstruction.

Previous work in fast neutron image reconstruction that would be adaptable to cone-beam effect removal has focused on computerized tomography utilizing either the Algebraic Reconstruction Technique (ART) [2, 3] or the Total Variation method (TV) [4, 5].

While ART can also be used to deconvolve the cone beam effect from an uncollimated point source, we examined a variation of the maximum entropy method (MEM) in our computational inversion because prior work by [6, 7] showed that MEM offers a qualitative advantage over ART when dealing with noisy images. MEM also offers an additional advantage over ART in that it minimizes artifacts by explicitly making the reconstructed image as noiseless as possible [8, 9].

On MEM vs TV, the comparision was not as clear cut, and there has not been scholarly work quantifying the advantages of either algorithm over the other. Past work in the TV by [5] showed how TV can be used to reconstruct and denoise any data and image linked through a linear operator. Nevertheless, we chose to implement MEM over TV in our work because MEM allowed us to easily define a Bayesian prior over the space of possible reconstructed images (see Appendix B) while it is not obvious how one would do so with TV. Adding a Bayesian prior to the reconstruction improves quality because it changes the default image from uniformly flat to the specified prior. For a spherical source which does not impart uniform neutron

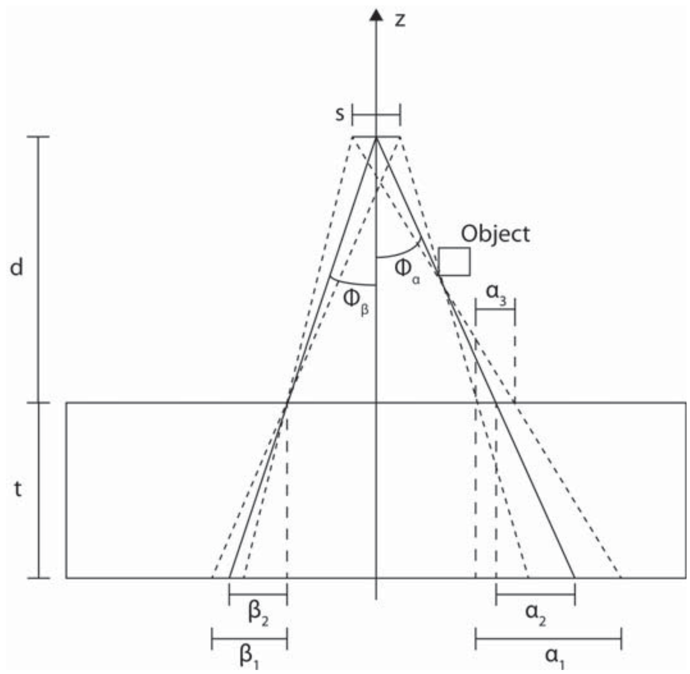

Figure 1: Cone beam effect for a feature focused at the detector's surface and for a feature focused away from the detector's surface

flux over the entire scintillator, the introduction of the correct prior reduces artifacts and improves contrast. This improvement is especially prominent in the peripherals of the image.

In this paper we describe our MEM based software for removing the cone beam effect, and apply the algorithm to a simulated fast neutron radiography image as a proof of concept. The algorithm is able to remove the part of the cone beam effect in the simulated image caused by the thick scintillator. With the help of the proposed algorithm, the reconstructed image appears to be taken with an infinitesimally thin scintillator (ITS). To the authors' knowledge this is the first algorithm presented which targets radiography image problems caused by thick scintillators.

We have arranged the paper as follows: Section 3 gives a reformulation of the cone beam problem, and Section 4 presents an overview of our algorithm. Section 5 details the MCNP simulation we used to generate our test images and Section 6 presents key algorithm details. Finally, we close with numerical results in Section 7, and conclusions along with future work in Section 8.

3 Cone beam effect

The cone beam effect (CBE) is a generalization of the concept of geometric unsharpness. The effect is dependent on the thickness of the radiation detector used, and the angle formed by the source to detector ray. When the source to detector distance is within an order of magnitude compared to the detector thickness, CBE becomes the prominent factor in image degradation .

As shown in Figure 1, when source to detector distance, d, and detector thickness, t, are fixed, the CBE can be completely characterized by Φ, the angle between the ray passing through the entry point and the z axis. Since this angular dependence is rotationally invariant about the z axis, any line from the source to a point on the surface of the detector can be rotated about the z axis to create a cone which is subject to the same blurring effect. It is this conical symmetry which gives this particular geometric unsharpness its name, the cone beam effect.

There are two ways to mitigate the CBE given a fixed d. The first is to reduce the source cross sectional

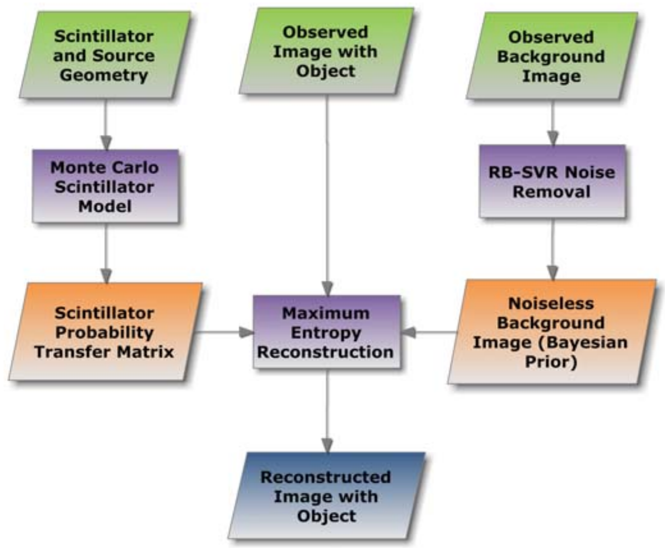

Figure 2: Algorithm Flowchart

area and for a feature focused at the surface of the detector, we can see this will at most reduce the blur from β_1 to β_2. The other approach is to reduce the thickness of the detector and from Figure 1, we see that this approach will yield an infinitely sharp point, when the detector becomes infinitesimally thin. For a feature not focused at the surface of the detector, taking $d \to 0$ will not recover a completely sharp image, α_3, but the resulting image quality is still superior to the approach of taking $s \to 0$, α_2, Figure 1. Thus, we focus our efforts on removing the primary effect of CBE and create a post processing routine to recover the image taken with an ITS when given an image taken with a detector of finite thickness t.

4 Algorithm overview

Our strategy for removing the cone beam effect from the uncollimated images comprises of MEM in conjunction with two supporting subroutines. MEM requires an input of a zero information image for noise suppression purposes and we accomplish this by modeling the ideal background image. Also, MEM requires a linear operator linking the ITS image to the observed image and we compute this operator through a Monte Carlo simulation of the neutron scintillator. Finally, we input the thick scintillator image with both the ideal background image and the linear scintillator model to reconstruct the ITS image.

5 Algorithm Test Procedures

For validation, we tested the algorithm with a typical fast neutron radiography problem that involves a low Z material hidden behind a high Z shield. We used MCNP simulation to generate both the ITS image and the thick scintillator image since MCNP can predict scattering by the high Z shield, the low Z object, and also scattering within the scintillator itself. To allow for future validation, we used realistic geometries and materials which can be readily replicated in a lab setting, Figure 3.

The simulation setup consists of a Cf-252 source imaging a composite test object placed behind a 1 inch thick lead shield. The test object contained plastics of different densities as well as metal features, Figure 3b. We purposely kept the source to detector distance small to ensure prominent CBE on the simulated radiographs.

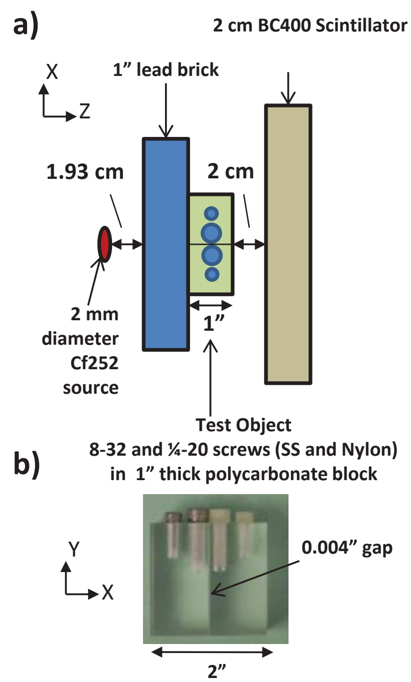

Figure 3: A) Top down view of MCNP model B) Expanded view of test object

$$P(e_n = x) = \frac{1}{3.37823} e^{(-x/1.025)} sinh(2.96x) \tag{1}$$

We ran two MCNP simulations, one for the thick scintillator and one for the ITS. In both runs, we used the neutron distribution shown in Eqn. 1, [10, 11] and a 2 mm diameter Cf-252 source. For the 2 cm thick detector, a 500 x 500 mesh heating tally was imposed on a Bicron BC 400 scintillator with the assumption that the heating tally directly translates to the light output of the scintillator. We used this MCNP run as the observed image input for our algorithm, Figure 4a. We also reran the simulation with the lead shield and test object removed to get a background image, Figure 4c. For the thick scintillator 1e11 Monte-Carlo neutrons confined to emission angles less than 40 degrees off the z-axis were used for each run.

For the ITS, a 500 x 500 radiography tally with MCNP's hybrid point detector model, which returns the incident neutron energy flux at each pixel, was used. This resulted in an essentially noiseless image which is shown in Figure 4b. As this model is essentially noiseless, it can be understood as the ideal detector image when the number of source particles is taken to infinity. With these simulated data, we can test our algorithm by inputting the observed image and seeing how well it reconstructs the ITS image.

6 Algorithm Implementation

In this section, we describe the maximum entropy method along with its two subroutines in detail.

6.1 Maximum Entropy Method

The problem of going from an observed image to an ITS image can be posed as a linear inversion. First we order the pixels in the observed image (size $M \times N$) from 1 to MN and reformulate the image into a vector based on the ordering. Next, we assume there exists a linear operator relating x, the $(MN \times 1)$ vector which represents the ITS image, and y, the observed image. We term this linear operator, A. Finally, we are left with a simple linear model relating the ITS image to the observed image, Eqn. 2.

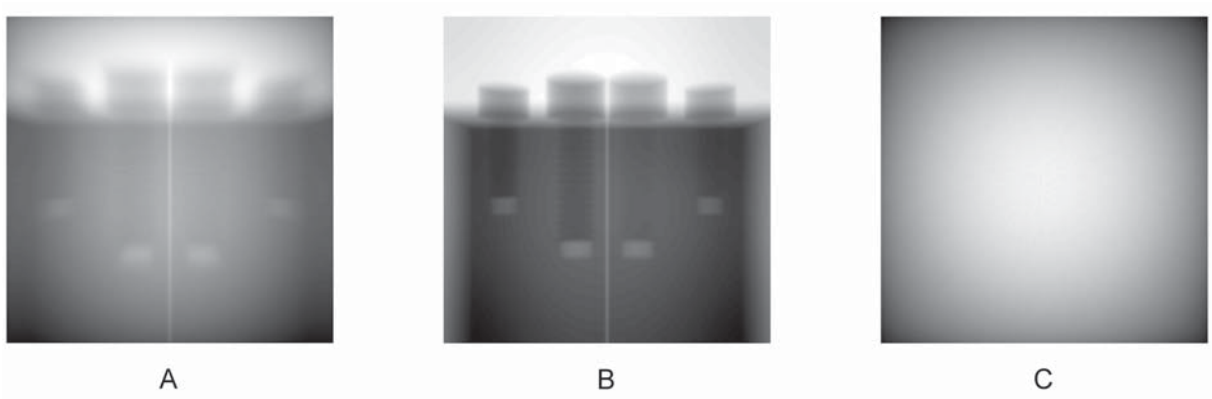

Figure 4: A) MCNP image with a 2cm thick scintillator B) MCNP image with the ideal thin scintillator C) MCNP background image

$$Ax = y \tag{2}$$

While $Ax = y$ can be solved via Gaussian-Jordan elimination, in practice we are faced with a noised perturbed version of the original equation, Eqn. 3.

$$Ax = y + \epsilon \text{ where } \epsilon \sim N(0, \sigma \cdot \ I_{MN}) \tag{3}$$

Given the additive noise ϵ, Eqn. 3 is ill posed and requires regularization for a viable solution. Many regularization techniques, such as ridge regression or the $L1$ loss, have been proposed for this ill posed problem but most lack a sound theoretical basis, [8, 12]. Out of these regularization techniques, the maximum entropy method (MEM) stands out because it allows the input of background image or Bayesian prior. MEM uses this Bayesian prior and the observed image to select a reconstructed image, x, which is most similar to the Bayesian prior while remaining statistically alike to the observed image when transformed with A. To accomplish this, MEM requires three inputs, A, y, σ and b, the zero information image. In most imaging applications of MEM, b is assumed to be the background image and here we do the same. We define b_i as the value of the background image for cell i. For numerical purposes, we normalized the background image such that $\sum_i b_i = 1$. Additionally, σ is a tuning parameter trading off noise suppression versus accuracy in the reconstructed image.

Once the inputs are defined, MEM approximates a solution to Eqn. 3 by solving Eqn. 4, which is always guaranteed to be well posed.

$$argmax\ S(x_1, \ldots, x_{MN}) = -\sum_i p_i \log(\tfrac{p_i}{b_i})$$

s.t.

$$\sum_{i=1}^{MN} \frac{(y_i - \sum_i A_i j x_j)^2}{\sigma} \leq MN \tag{4}$$

Where

$$p_i = \frac{x_i}{\sum_j x_j}$$

Eqn. 4 states that the optimal reconstruction is found by maximizing the Shannon entropy of the reconstruction while making sure that the reconstructed image, when operated on by A, is still statistically similar to the observed image. Intuitively, maximizing the Shannon entropy flattens the reconstruction and grants MEM its noise suppression characteristics. Also, the statistical similarity criterion forces the

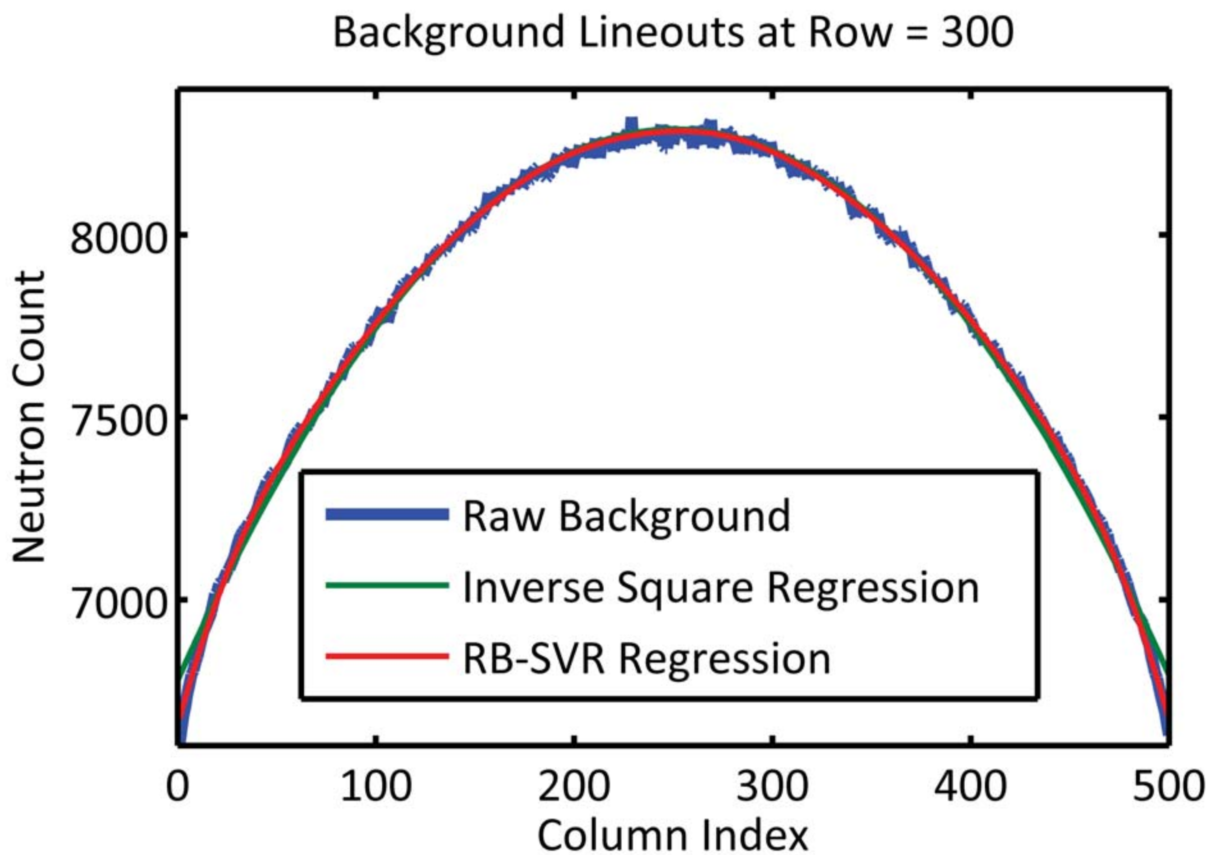

Figure 5: Comparison of the RB-SVR noise removal versus inverse square law noise removal

reconstruction to approximately satisfy Eqn. 3 and is what enables MEM to invert A. A derivation of MEM, is given in Appendix B.

6.2 Bayesian Prior Subroutine

Since maximum entropy reconstruction depends heavily on the Bayesian prior to flatten the image, it is crucial that we have an intensity map of the scintillator response for when there is no object of interest. This map also needs to be as smooth as possible as any noise in the Bayesian prior will be magnified in the reconstructed image. We start by noting that the neutron flux is only quasi-radial symmetric because the CF-252 source is not spherical and is large enough to make point particle approximations insufficient.

While we can generate a noise-free background image through a Monte-Carlo routine, the amount of simulated particles required for smooth convergence will require many computer days. Also, this approach assumes that the simulation geometry and experimental geometry are one and the same. Any geometrical artifacts from incomplete calibration will result in errors propagating through the reconstruction. As we expect calibration errors to happen when we eventually deploy this technique, we searched for techniques based on smoothing an observed background image. The top two candidates were inverse square regression and residual boosted support vector regression (RB-SVR). Previously work with SVR in image processing, [13, 14], showed great success but we found RB-SVR demonstrated better empirical performance, Figure 5. Thus, we choose RB-SVR as our smoothing algorithm. Rather than writing our version of RB-SVR, we wrapped LIBSVMs support vector regression algorithm, [15], in our own residual boosting module. A derivation of RB-SVR is given in Appendix A.

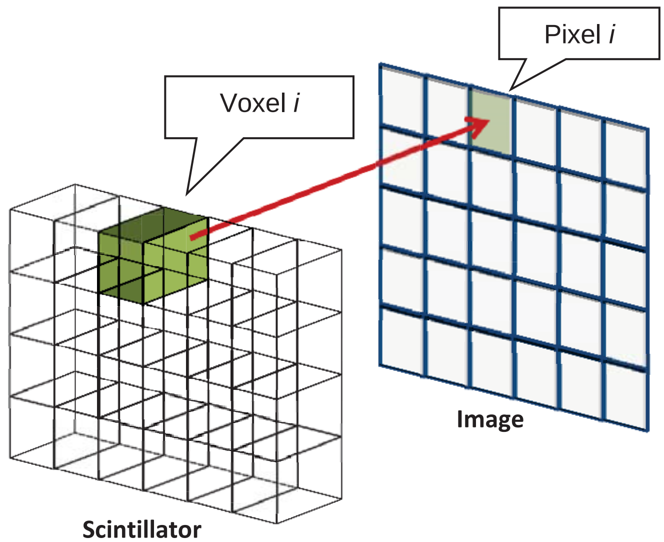

Figure 6: Graphical demonstration of the bijection between voxel and pixel

6.3 Linear Scintillator Model

For our linear scintillator model, we opted for a full discrete treatment instead of any continuous approximation. First we partitioned the scintillator into voxels, volumetric pixels. Next, we assumed that the distance to first interaction of a neutron in the scintillator is an exponential random variable with its mean equal to the mean free path of the neutron. Finally, we assumed that all neutrons give up all their energy on the first interaction so there is no scattering within the scintillator.

Working off the above assumptions, we defined a bijection between each pixel in the observed image and a voxel on the scintillator, Figure 6.

If any neutron interacts with the scintillator in voxel i, we assume pixel i increases its intensity count by a constant factor. The prior assumption is justified because each voxel interacts with enough neutrons to ensure central limit convergence, and the ratio of standard deviation of neutron count to mean neutron count is less than 0.05.

In the framework of the discrete voxel scintillator, we wish to solve the number of neutrons incident on each voxel given the number of neutrons terminating in each voxel. Neglecting the intensity variance, the number of incident neutrons is exactly the response of the ideal infinitesimally thin detector. Thus removing the cone-beam effect is equivalent to solving for the number of incident neutrons.

6.4 Probability Transfer Matrix

Given our voxel scintillator model, we now clarify our linear model, Eqn. 2. We redefine x as the vector consisting of the incident counts for each voxel, and y as the vector consisting of the termination counts for each voxel. Now A becomes the probability transfer matrix (PTM) between incident voxel and termination voxel, such that $[A]_{ij}$ denotes the probability a neutron incident on voxel i will terminate in voxel j, Figure 7.

While A can be determined analytically for certain neutron emission distributions and scintillator compositions, we follow the time tested approaches of [16, 17, 18, 19] and obtain A through Monte Carlo simulation, whose pseudocode is listed below.

for $i = 1 \to MN$ **do**
  for $n = 1 \to numNeutrons$ **do**
    Sample α_n, the entry point on voxel i's surface

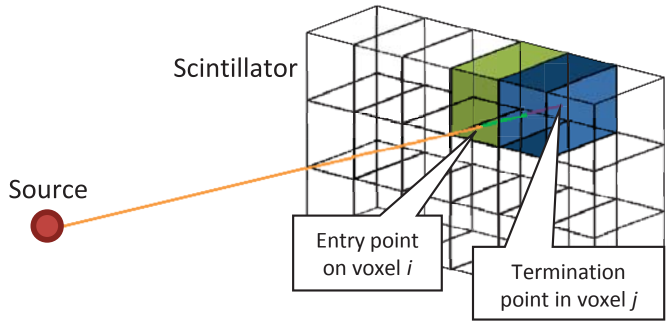

Figure 7: Neutron incident on voxel i contributing to light output in voxel j

Sample β_n, the emission point on the sources surface
Sample e_n, the neutron energy for neutron n
Calculate λ_n given e_n
Sample p_n, the penetration distance, given λ_n
Ray Trace from α_n to β_n and find psi_n, the termination point, given p_n
Calculate which voxel $j \ni {}_n$
Add $\frac{1}{numNeutrons}$to $[A]_{ij}$
end for
end for

For sampling the emission energy, e_n, we utilized Eqn. 1 and to calculate the mean free path given neutron energy, we utilized the Evaluated Nuclear Reaction Data library, [20], and found the mean free path of neutrons with energy ranging from 1e-10 to 20 MeV in Bicron BC-400 through its molecular formula. We show the mean free path in Figure 8. Technically, the energy distribution of the neutron incident to the scintillator after passing through the lead shield and the test phantom will not be the same as Eqn. 1 due to spectral hardening. However looking at Figure 8, we see that the mean free path of neutrons in BC-400 is roughly flat for energies of 2 MeV to 4 MeV. Since the mean neutron energy of Cf-252 is 2.314 MeV, we see that spectral hardening by the lead shield and the test phantom can shift the mean neutron energy upward by a factor of 2 and still have negligible effect on average mean free path.

Finally, as a check, we applied A to a uniformly flat image. Based on the inverse square law, we expect to see a concave sink extending out from center of the image after transformation, confirmed in Figure 9.

7 Numerical Results

First, to estimate the Bayesian prior of the scintillator response with no object, we used MCNP to simulate a background image as discussed in Section 4. Figure 10a shows a histogram of the neutron counts per pixel of Figure 4c which illustrates the noise in the background image which we have to minimize for the reconstruction. Figure 10b shows a histogram of the neutron counts per pixel after RB-SVRs estimate of the background image. The uneven landscape of the difference between the two demonstrates the successful noise removal of RB-SVR processing, Figure 10c.

Next, we generated two versions of the PTM in order to study the effects of Monte Carlo noise on the reconstruction. The first PTM had 2e6 neutrons per voxels and the second PTM had 2e7 neutrons per voxel. Given a desktop Intel I7 950, we were able to simulate 2e6 neutrons per second and build a PTM matrix with numNeutrons set at 2e7 in 140 CPU hours.

Finally, after combining the RB-SVR Bayesian prior and the observed image, we used MEM to reconstruct our estimate of the original image, shown in Figure 11b and Figure 11c. The original observed image, Figure 4a, and the ideal image, Figure 4b, are reproduced and enlarged in Figure 11a and 11d for convenience and

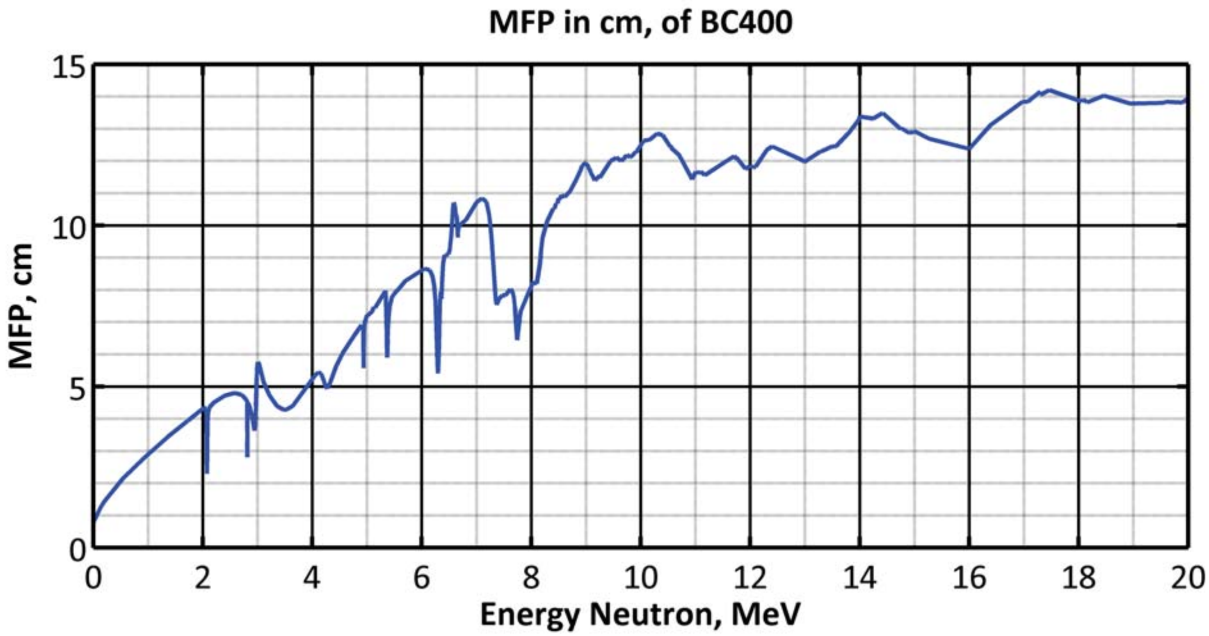

Figure 8: Mean free path vs neutron energy for neutrons entering Bicron BC-400

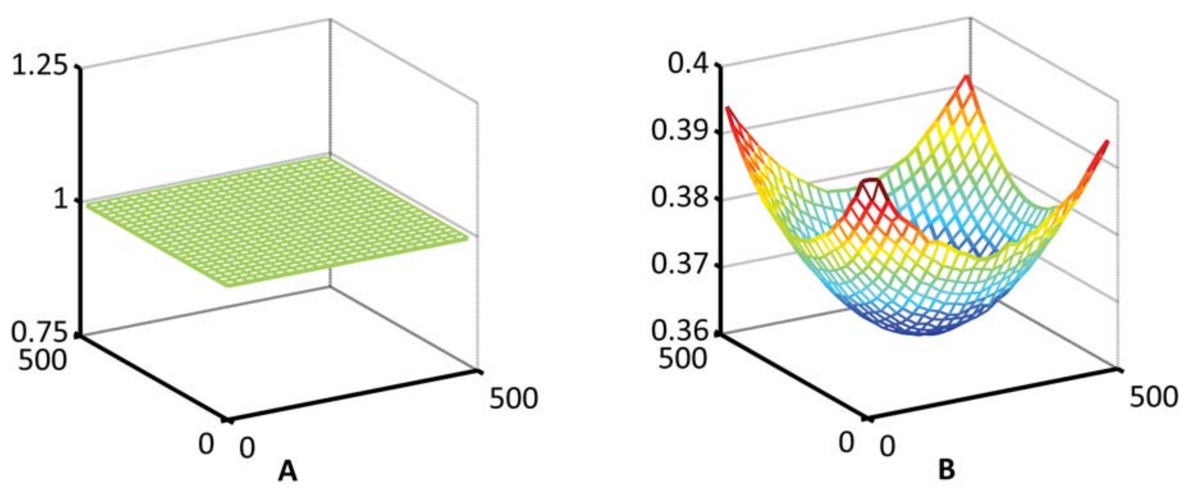

Figure 9: A uniformly flat image before (A) and after (B) applying our linear scintillator model

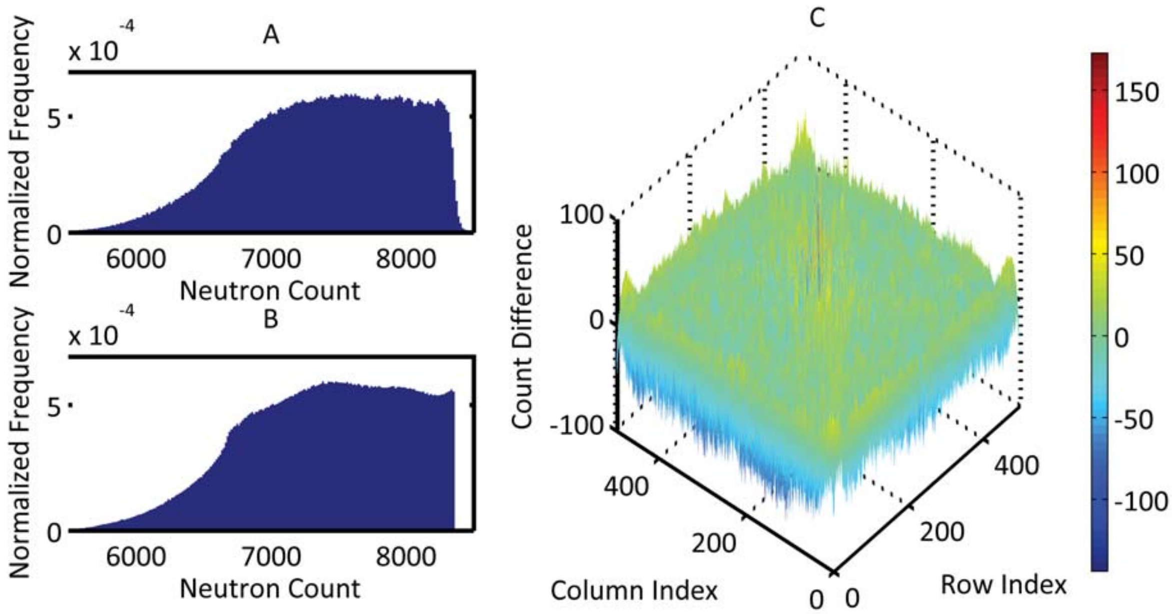

Figure 10: A) Histogram of MCNP background image's neutron counts per pixel. B) Histogram of background image's neutron counts per pixel after RB-SVM processing. C) Difference in neutron counts per pixel

ease of comparison. On the same desktop, MEM reconstruction took 20 seconds per image.

For both reconstructions, we observed suppression of CBE and restoration of hard edges. Comparing Figure 11b and Figure 11c, we see that Monte Carlo noise in the probability transfer matrix carries through to image reconstruction. The improvement in reconstruction quality between Figure 11b and Figure 11c, is due to the use of an order of magnitude more simulated particles in the construction of Figure 11cs PTM.

In this formulation, the reconstructed image is completely defined by the A, b, and σ. We would like to stress that the reconstruction is defined in terms of a global optimzation problem and with A, b, and σ fixed, the reconstruction is also completely independent of any initial conditions to the MEM problem. In an effort to explore the robustness of the algorithm to noise, we added increasing levels of Gaussian white noise to Figure 4a before removing the CBE with our algorithm. For each level of Gaussian noise, we set the reconstruction parameter σ equal to the σ of the Gaussian noise. We measured the degradation effects of the white noise by computing the normalized RMSE between the reconstruction with noise and the reconstruction without noise. We chose normalized RMSE because of it's natural interpretation as the average percentage difference between pixel values and define it below.

$$\mathrm{Norm}(x)_i = \frac{x_i}{\max(x)}$$

$$\mathrm{RMSE}_{NORM} = \sqrt{\frac{\sum_i^N \left(\mathrm{Norm}(\hat{x})_i - \mathrm{Norm}(x)_i\right)^2}{N}}$$

Next as an objective reference, we estimated the signal to noise ratio (SNR) for the original image and its corrupted copies. For the estimation, we chose a 50×50 pixel background area in the image, Figure 4a, and calculated the mean well as the standard deviation of the neutron counts in the area.

As shown in Figure 12, the relationship between RMSE and SNR is approximately linear. Assuming a Poisson emission model for the neutron counts, white noise at the $\sigma = 50$ level increases variance per pixel by 2500 or 33% (from the average neutron count of 7500). However, this 33% increase in variance only results

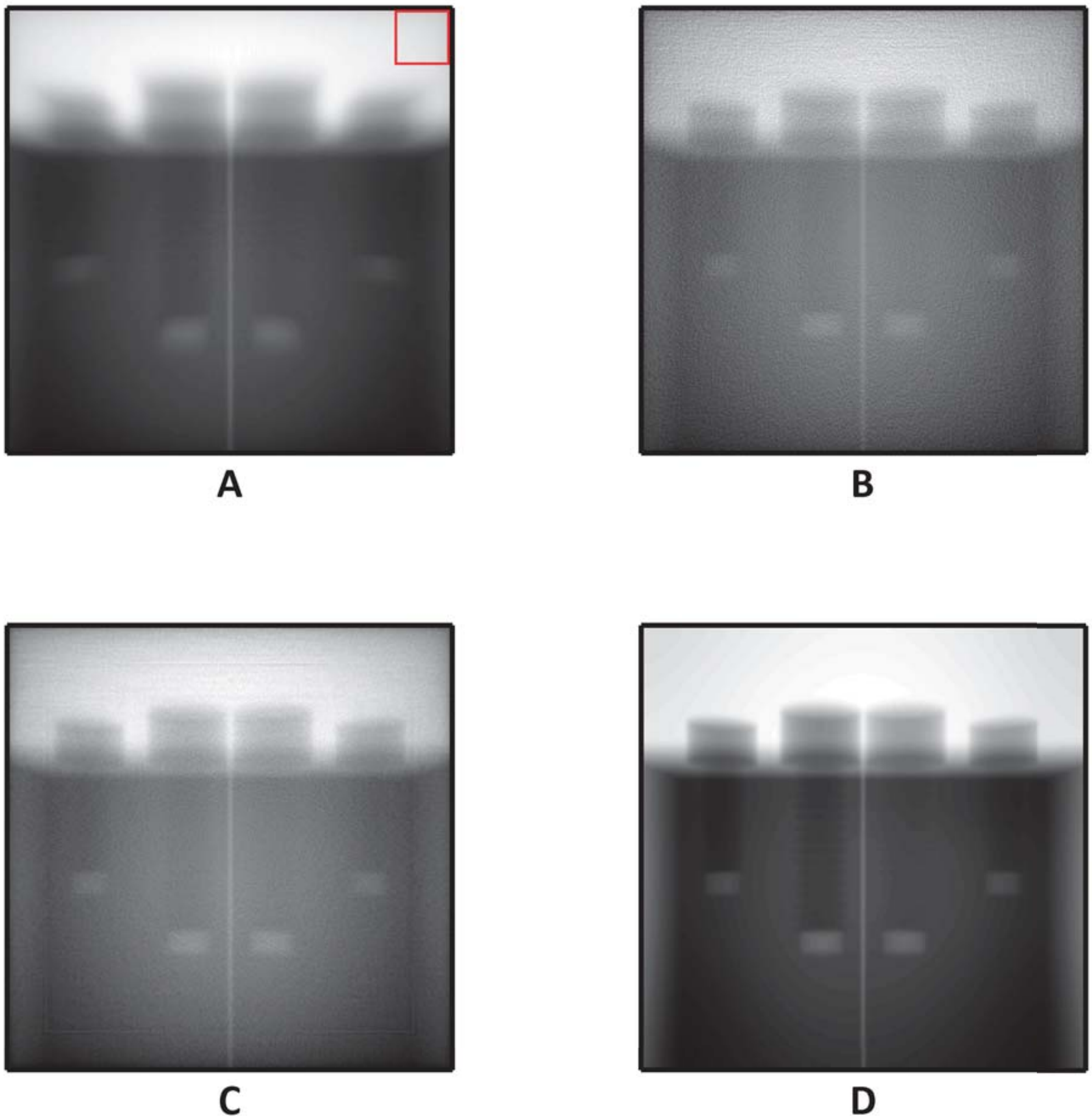

Figure 11: A) MCNP simulated image with finite thick scintillator. B) Restored image with a PTM built with 2e6 neutrons per voxel. C) Restored image with a PTM built with 2.5e7 neutrons per voxel. D) MCNP simulated image with ideal thin scintillator. NOTE: SNR values in Figure 12 are estimated from pixels within the red square in (A).

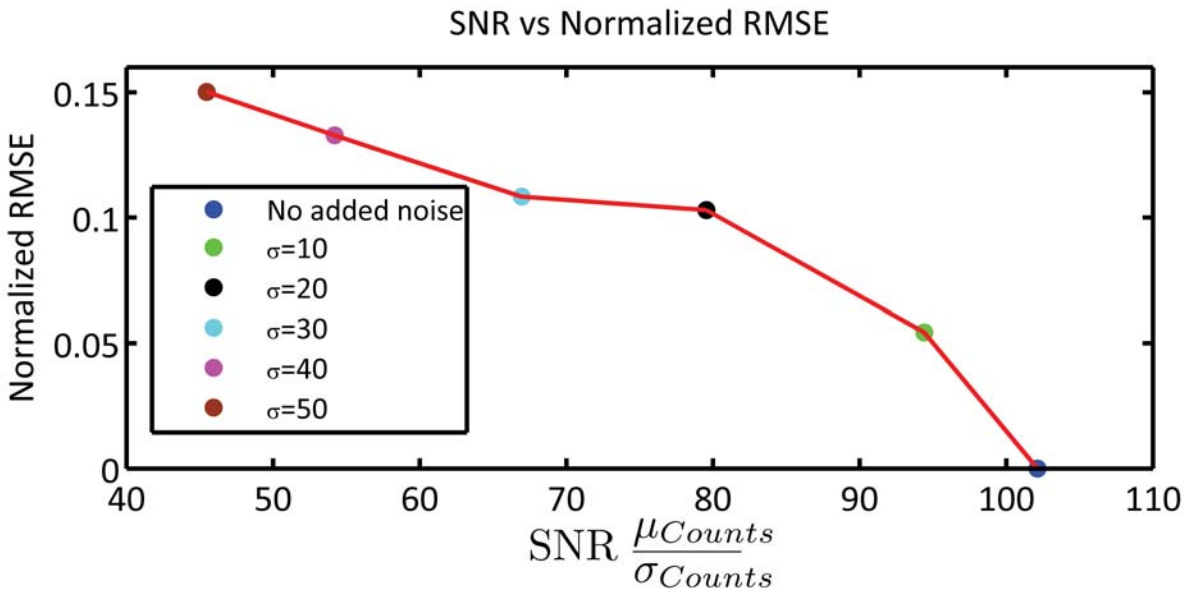

Figure 12: RMSE vs SNR

in a 15% difference in reconstruction value, Figure 12, demonstrating that the algorithm is robust to modest amounts of noise.

8 Conclusion

It has been shown by simulation that the cone beam effect can be practically removed from an uncollimated fast neutron image through a simple model of the scintillator response and source to scintillator geometry.

Residual boosted support vector regression was used to smooth the background intensity and large scale Monte Carlo simulation was used to generate a linear approximation of the scintillator response to a near field divergent neutron source. Finally, the maximum entropy method was used to invert the scintillator response from an MCNP simulated observed image.

Overall, the discussed reconstruction techniques could reduce exposure times or required source intensity without undesirable object blurring on the image by both allowing closer source-to-detector distances to maximize incident radiation flux and the use of thicker scintillators with higher efficiencies. In addition to neutron imaging the technique should also be applicable, with the right PTM, for high energy gamma or x-ray radiography using thick scintillators.

Future work will revolve around calibrating an experimental setup which mimics our simulation geometry and applying the algorithm to an empirical image. Also, while we chose to utilize MEM as the current inversion algorithm because of the ease with which it can accomodate a Bayesian prior, a systematic comparison which will quantify the advantages and disadvantages of the method still needs to be done between MEM and the other two contending algorithms, ART and TV. That comparision, however, is non-trivial due to the complexities involved with adapting and implementing ART and TV to the CBE problem. In addition, Figure 11 demonstrated the importance of quality of the PTM. To the authors' best knowledge, the only way to improve PTM quality is to utilize more particles in its construction. Since the Monte-Carlo routine used to generate the PTM is parallelizable, other future work will revolve on adapting the software to run on a high performance computing cluster to quickly generate high quality PTM matrices.

9 Acknowledgements

We thank Jim Hall, Brian Rusnak, and Phil Kerr at LLNL for discussions and guidance on neutron imaging as well as the loan of imaging equipment and facilities. This work was supported in part by the US. Department of Energy Na-22 Office of Nonproliferation Research and Development under the Radiological Source Replacement program and was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

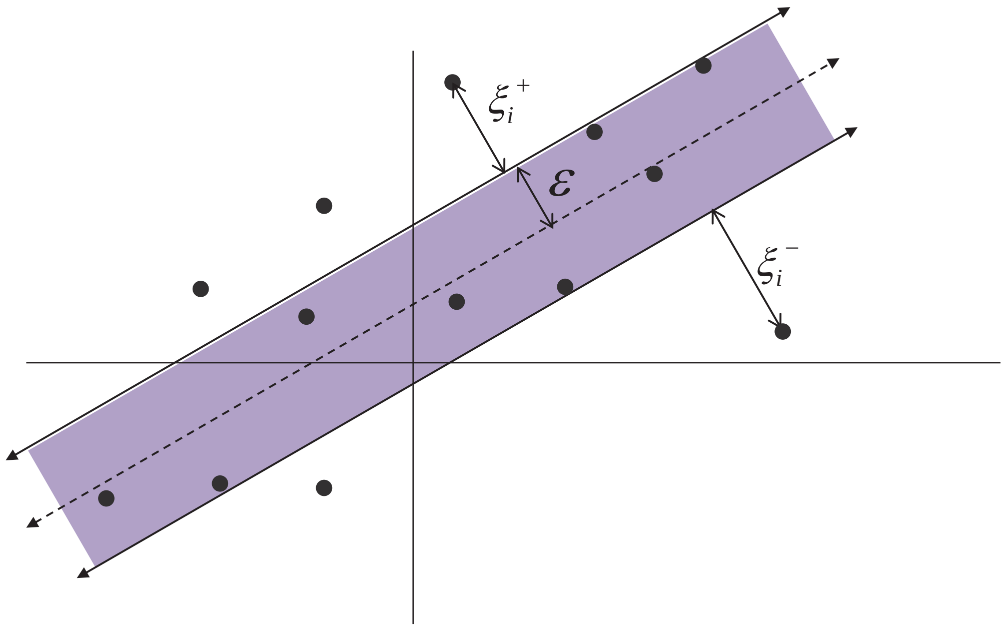

Figure 13: A visualization of a SVR regression in two variables

A Support Vector Regression Theory

Developed by [21] in 1990, support vector regression (SVR) is a machine learning technique which can approximate nonlinear functions, [22].

Given a training set, $T : (x_1, y_1), \ldots, (x_N, y_N) \in R^m \times R$, we wish to approximate a function $f(x_i) \approx y_i$ s.t.

$$f(x) =< \phi(w), \phi(x_i) > +b \tag{5}$$

Where $< \cdot, \cdot >$ denotes the inner product and $\phi(x_i)$ is a nonlinear mapping from R^m to a higher dimensional space. The parameters w, and b are solved via minimization of the following cost function, R_f.

$$\text{argmin } R_f(w, b) = \frac{1}{2}\|w\|^2 + CR_{emp} \tag{6}$$

Here, R_{emp} measures empirical risk, $\|w\|^2$ measures model complexity and C is a regularization parameter which balances model complexity and training set performance. We define R_{emp} as

$$R_{emp} = \frac{1}{N}\sum_{i}^{N} |y_i - f(x_i)|_\epsilon \tag{7}$$

where $| \cdot |_\epsilon$, termed by [21] as the ϵ-insensitive loss is defined as

$$|y_i - f(x_i)|_\epsilon = \max\{0, |y_i - f(x_i)| - \epsilon\} \tag{8}$$

Thus as a result of Eqn. 8, regression estimates which err by less than do not factor into the cost function resulting in an insensitive tube around the regression estimates, Figure 13.

We wish to reformulate Eqn. 6 as a quadratic programming problem for tractable computation so we introduce slack variables ξ_i^+, and ξ_i^-. The two slack variables, ξ_i^+ and ξ_i^-, measures the deviation of observation i above and below the surface of the ϵ tube respectively. This formulation is termed the ϵ-SVR by [21].

$$
\begin{gathered}
\text{argmin } F(w,b,\xi^-,\xi^+) = \tfrac{1}{2}\|w\|^2 + \tfrac{C}{N}\textstyle\sum_i(\xi_i^+ + \xi_i^-) \\
\text{Subject to} \\
y_i - f(x_i) \le \xi_i^+ + \epsilon \\
f(x_i) - y_i \le \epsilon + \xi_i^- \\
\xi_i^-, \xi_i^+ \ge 0, i = 1, \ldots, N
\end{gathered} \tag{9}
$$

While we now have a well posed quadratic programming problem, we are required to set the parameter ϵ a priori. This is unsatisfactory because ϵ is highly data dependent and can range over $[0, \infty)$. To remove the burden of selecting ϵ , [23] introduced the ν-SVR. ν-SVR introduces a new parameter ν and redefines the optimization problem as follows.

$$
\begin{gathered}
\text{argmin } F(w,b,\xi^-,\xi^+,\epsilon) = \tfrac{1}{2}\|w\|^2 + C[\nu\epsilon + \tfrac{1}{N}\textstyle\sum_i(\xi_i^+ + \xi_i^-)] \\
\text{Subject to} \\
y_i - f(x_i) \le \xi_i^+ + \epsilon \\
f(x_i) - y_i \le \epsilon + \xi_i^- \\
\xi_i^-, \xi_i^+ \ge 0, i = \{1, \ldots, N\}
\end{gathered} \tag{10}
$$

Now, ϵ is a variable featured in the optimization problem and is no longer a parameter. While we have substituted ν for ϵ, ν is bounded $\in$ [0, 1] and has an intuitive meaning as the maximum fraction of y_i's allowed to err by more than ϵ.

A.1 Kernel Functions

Much of the SVRs advantage comes from its projection of data into higher dimensional space, $\phi(\cdot)$. Let $K(x_i, x_j) =< \phi(x_i), \phi(x_j) >$. $K(x_i, x_j)$ is called a kernel function and it provides the benefit of a high dimensional space without explicit computation. For example, the second order polynomial kernel $K_{P2}(x_i, x_j) = (x_i \cdot x_j)^2$ is equivalent to $< \phi(x_i), \phi(x_j) >$ with $\phi : R^2 \to R^3$ s.t.

$$
\phi \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x^2 \\ y^2 \\ xy \end{bmatrix}
$$

Kernels can also be chosen with a priori knowledge and since we know that the true background intensity map will be smooth, we choose a kernel function which favors smoothness.

$$
K(x_i, x_j) = e^{(-\gamma\|x_i - x_j\|^2)} \tag{11}
$$

This kernel is known as the Gaussian radial basis function (RBF) and the ϕ associated with this kernel projects the data into an infinite dimension Hilbert space. However, this kernel also introduces an addition parameter γ which must be optimized during training.

A.2 Residual Boosting

While ν-SVRs have great native performance, they can be combined through boosting for even better results. In residual boosting, a particular form of ensemble learning, the regression target is iteratively simplified so the machine learner can capture higher order effects in successive iteration.

Residual boosting accomplishes this by defining $t_{i,k}$, the regression target for observation i at iteration k as

$$
t_{i,k} = \begin{cases} y_i - \sum_{m=0}^{k-1} f_m(x_i) & \text{if } k > 0 \\ y_i & \text{if } k = 0 \end{cases} \tag{12}
$$

Thus, the k^{th} iteration of the machine learner only tries to capture the residuals of the prior $k-1$ iterations. For the final estimate, we take sum of all the regression functions, Eqn. 13.

$$\hat{f}(x_i) = \sum_{n=0}^{K} f_n(x_i) \tag{13}$$

K, the maximum number of boosting iterations, is decided in advance and we found that regression accuracy converge for $K > 3$.

B Maximum Entropy Method

Developed in 1984 by [8], MEM is a image processing technique well documented in the astronomy community. The original algorithm was developed to deconvolute atmospheric point spread functions from telescope images but the technique has been generalized to a variety of fields due to the algorithms ability to invert any linear operator, [24, 25, 26, 27, 28]. As an added bonus, MEM also reconstructs the flattest image possible given the observed data, reducing the number of post reconstruction artifacts [8, 9].

To derive the MEM formulation, we assume that we have K balls (neutrons) and when thrown, each ball is independent and is equally likely to land in any of the MN buckets (voxels). We do not know the actual distribution of the balls among the buckets but the best guess would be the distribution with the highest probability. Since each particular distribution is a realization of a multinomial random variable, we can find the most probable distribution by maximizing the probability of a certain distribution happening.

$$\text{argmax } P(x_1, \ldots, x_{MN}) = \frac{K!}{\prod_{i=1}^{MN} x_i!} \frac{1}{MN}^{-K} \tag{14}$$

Maximizing $P(x_1, \ldots, x_{MN})$ is equivalent to maximizing any monotonic transform of $P(x_1, \ldots, x_{MN})$, so we choose to maximize

$$F(x_1, \ldots, x_{MN}) = \frac{1}{K} \log\left(P(x_1, \ldots, x_{MN})\right) - \frac{1}{K} \log\left(\frac{1}{MN}^{-K}\right) \tag{15}$$

$$= \frac{1}{K} \log\left(\frac{K!}{\prod x_i}\right)$$

$$= \frac{1}{K}\left[\log\,(K!) - \sum_{i=1}^{MN} \log\,(x_i!)\right] \tag{16}$$

Since $K \sim 1e18$, we can use Sterlings approximation, $\log(K!) \approx K\log(K) - K$, on Eqn. 16.

$$= \frac{1}{K}\left[K\log\,(K) - \sum x_i \log\, x_i - K + \sum x_i\right]$$

$$= \log\,(K) - \sum \frac{x_i}{K} \log\left(\frac{x_i}{K} \cdot K\right)$$

$$= \log\,(K) - \sum \frac{x_i}{K} \log\,(K) - \sum \frac{x_i}{K} \log\left(\frac{x_i}{K}\right)$$

$$= \left(1 - \sum \frac{x_i}{K}\right) \log(K) - \sum \frac{x_i}{K} \log\left(\frac{x_i}{K}\right)$$

$$= -\sum \frac{x_i}{K} \log\left(\frac{x_i}{K}\right)$$

$$= -\sum p_i \log\,(p_i) \tag{17}$$

Looking at Eqn. 17 we see that it is equivalent to the Shannon entropy of a multinomial distribution, Eqn. 18.

$$S_{Shannon} = \sum_{i=1}^{MN} p_i \log(p_i) \tag{18}$$

Thus we see maximizing the image entropy is equivalent to finding the most probable image. However, we cannot blindly apply the balls and buckets model to our problem as each voxel possesses a different solid angle area, and thus receive different amounts of neutron flux, we modify each p_i in Eqn. 18 with a Bayesian prior, b_i, to correct for the neutron flux difference, Eqn. 19.

$$= -\sum p_i \log\left(\frac{p_i}{b_i}\right) \tag{19}$$

Looking at Eqn. 19, we see that unconstrained optimization will simply return the b_is as the reconstructed image. Thus, we add the Chi-Square constraint, Eqn. 20, to guarantee the reconstructed image is statistically similar to the observed image when linked through the linear operator. For a more complete discussion on the Chi-Square test and its assumptions, please see [29].

$$\sum_{i=1}^{MN} \frac{(y_i - \sum[A]_{ij}x_j)^2}{\sigma} \leq MN \tag{20}$$

Combining Eqn. 19 and Eqn. 20, we arrive at the full formulation of MEM.

$$\begin{aligned} & \text{argmax} \;\; -\sum p_i \log\left(\tfrac{p_i}{b_i}\right) \\ \text{s.t.} & \\ & \textstyle\sum_{i=1}^{MN} \frac{(y_i - \sum[A]_{ij}x_j)^2}{\sigma} \leq MN \end{aligned} \tag{21}$$

To solve Eqn. 21, a large scale convex optimization problem involving MN variables, iterative preconditioned gradient descent or quasi-Newton methods can be used. For brevity, we will not restate the various numerical algorithms but refer the reader to [8, 30].

References

[1] J. Hall, B. Rusnak, and P. Fitsos, “High-energy neutron imaging development at llnl in 8th world conference on neutron radiography,” No. 8th, (Gaithersburg, MD.), 2006.

[2] R. Gordon, R. Bender, and G. T. Herman, “Algebraic reconstruction techniques (art) for 3-dimensional electron microscopy and x-ray photography,” *Journal of Theoretical Biology*, vol. 29, no. 3, p. 471, 1970.

[3] J. Guo, T. Bucherl, Y. Zou, and Z. Guo, “Study on beam geometry and image reconstruction algorithm in fast neutron computerized tomography at nectar facility,” *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment*, vol. 651, no. 1, pp. 180–186, 2011.

[4] S. Osher, M. Burger, D. Goldfarb, J. Xu, and W. Yin, “An iterative regularization method for total variation-based image restoration,” *Simul*, vol. 4, pp. 460–489, 2005.

[5] E. Y. Sidky and X. Pan, “Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization,” *Physics in Medicine and Biology*, vol. 53, no. 17, p. 4777, 2008.

[6] G. T. Herman, “Mathematical optimization versus practical performance - a case-study based on the maximum-entropy criterion in image-reconstruction,” *Mathematical Programming Study*, vol. 20, pp. 96–112, 1982.

[7] K. J. Myers and K. M. Hanson, “Comparison of the algebraic reconstruction technique with the maximum entropy reconstruction technique for a variety of detection tasks,” *Medical Imaging IV: Image Formation*, 1990.

[8] S. F. Gull and G. J. Daniell, “Image-reconstruction from incomplete and noisy data,” *Nature*, vol. 272, pp. 686–690, 1978.

[9] D. S. Sivia, “Introduction to maximum-entropy,” *Institute of Physics Conference Series*, pp. 245–256, 1989.

[10] J. Terrell, “Fission neutron spectra and nuclear temperatures,” *Physical Review*, vol. 113, pp. 527–541, 1959.

[11] MCNP X-5 Monte Carlo Team, “MCNP a general purpose monte carlo n-particle transport code,” tech. rep., Los Alamos National Laboratory, 2003.

[12] W. Q. Yang and P. Lihui, “Image reconstruction algorithms for electrical capacitance tomography,” *Measurement Science and Technology*, vol. 14, p. R1, 2003.

[13] H. Gomez-Moreno, S. Maldonado-Bascon, F. Lopez-Ferreras, and P. Gil-Jimenez, *Removal of impulse noise in images by means the use of support vector machines*, vol. 2687 of *Lecture Notes in Computer Science*, pp. 536–543. Berlin: Springer-Verlag Berlin, 2003.

[14] D. Xiaoying, Y. Dinghui, P. Jiming, G. Xin, and Y. Baojun, “Noise reduction and drift removal using least-squares support vector regression with the implicit bias term,” *Geophysics*, vol. 75, pp. V119–V127V127, 2010.

[15] C. Chang and C. Lin, *LIBSVM: a library for support vector machines*. 2001.

[16] P. P. Dunphy, E. L. Chupp, and C. M. Jensen, “Monte-carlo simulation of the response of nai scintillator to medium-energy neutrons,” *Nuclear Instruments & Methods in Physics Research Section a-Accelerators Spectrometers Detectors and Associated Equipment*, vol. 330, pp. 199–209, 1993.

[17] B. Gustafss and O. Aspelund, “Detec a subprogram for simulation of fast-neutron detection process in a hydro-carbonous plastic scintillator,” *Nuclear Instruments & Methods*, vol. 48, no. 1, p. 77, 1967.

[18] W. W. Lindstro and B. D. Anderson, “Improved computer simulation of detection process in ne-102 and ne-213 scintillators for 0 to 4.8 mev neutrons,” *Nuclear Instruments & Methods*, vol. 98, no. 3, p. 413, 1972.

[19] J. Poitou and C. Signarbi, “Monte-carlo simulation of capture and detection of neutrons with large liquid scintillators,” *Nuclear Instruments & Methods*, vol. 114, no. 1, pp. 113–119, 1974.

[20] A. Trkov, M. Herman, and D. Brown, “Endf-6 formats manual, data formats and procedures for the evaluated nuclear data file, endf/b-vi and endf/b-vii,” Tech. Rep. BNL-90365-2009, Brookhaven NL, 2011.

[21] C. Cortes and V. Vapnik, “Support-vector networks,” *Machine Learning*, vol. 20, pp. 273–297, 1995.

[22] A. J. Smola and B. Scholkopf, “A tutorial on support vector regression,” *Statistics and Computing*, vol. 14, pp. 199–222, 2004.

[23] B. Scholkopf, A. Smola, R. Williamson, and P. Barlett, “New support vector algorithms,” *Neural Computation*, 2000.

[24] T. Elfving, “An algorithm for maximum-entropy image-reconstruction from noisy data,” *Mathematical and Computer Modelling*, vol. 12, pp. 729–745, 1989.

[25] G. Minerbo, “Maximum entropy reconstruction from cone-beam projection data,” *Computers in Biology and Medicine*, vol. 9, pp. 29–37, 1979.

[26] P. Schleger, A. Puig-Molina, E. Ressouche, O. Rutty, and J. Schweizer, “A general maximum-entropy method for model-free reconstructions of magnetization densities from polarized neutron diffraction data,” *Acta Crystallographica Section A*, vol. 53, pp. 426–435, 1997.

[27] H. Shioya and K. Gohara, “Maximum entropy method for diffractive imaging,” *Journal of the Optical Society of America a-Optics Image Science and Vision*, vol. 25, pp. 2846–2850, 2008.

[28] D. Tatchev and R. Kranold, “Maximum-entropy method as a routine tool for determination of particle size distributions by small-angle scattering,” *Journal of Applied Crystallography*, vol. 37, pp. 32–39, 2004.

[29] P. E. Greenwood and M. S. Nikulin, *A Guide to Chi-Squared Testing*. Wiley, 1996.

[30] J. Myrheim and H. Rue, “New algorithms for maximum entropy image restoration,” *CVGIP: Graphical Models and Image Processing*, vol. 54, pp. 223–238, 1992.