

Inference for a Proton Accelerator Using Convolution Models

H. Lee*, B. Sansó*, W. Zhou*, and D. Higdon†
[herbie,bruno,zhouwn]@ams.ucsc.edu, dhigdon@lanl.gov

Abstract

Proton beams present difficulties in analysis because of the limited data that can be collected. The study of such beams must depend upon complex computer simulators that incorporate detailed physical equations. The statistical problem of interest is to infer the initial state of the beam from the limited data collected as the beam passes through a series of focusing magnets. We are thus faced with a classic inverse problem where the computer simulator links the initial state to the observables. We propose a new model for the initial distribution which is derived from the discretized process convolution approach. This model provides a computationally tractable method for this highly challenging problem. Taking a Bayesian perspective allows better estimation of the uncertainty and propagation of this uncertainty to predictions.

Keywords: computer simulator, inverse problem, density estimation, Bayesian statistics

1 Introduction

Particle accelerators are a major tool in modern experimental physics. In order to be an effective tool, it is necessary for scientists to be able to understand the characteristics of the particle beam and to optimally focus the beam. The combination of physical laws and fast computers allows an accurate description of the beam if the initial conditions are known. However, these conditions cannot be directly measured, and thus inferring these conditions

*University of California, Santa Cruz

†Los Alamos National Laboratory

is the primary statistical goal. In our work we look at a particular proton accelerator, but the methodology developed herein should easily generalize to other accelerators (as well as other applications) and can play an important role in the design of future accelerators.

This problem is difficult because of the limited information that can be collected about the particle beam. What is measured during an experiment are one-dimensional histograms of the relative frequencies of the particles at certain points along the path of the beam. From this information, scientists need to learn about the initial distribution of the particles in both their x -dimension and y -dimension phase spaces, i.e., their position and momentum in each dimension. A key computational tool is a detailed computer simulator which predicts the evolution of the beam given a particular set of initial conditions. Thus we have a classic inverse problem, where we are attempting to learn about the unobservable initial state from highly transformed and simplified data, with computer code providing the link (see for example, Yeh, 1986). Inference involves an iterative procedure whereby initial states are proposed, run through the simulator, the predicted results computed and compared to the observed data, and then the initial proposal is modified in an attempt to improve the match between the computed results and the data.

We model the initial distributions nonparametrically via a convolution of a discrete gamma process with a Gaussian kernel. This approach turns out to be equivalent to using a mixture of Gaussians for density estimation, except that the mixture components are fixed to be identical (like the case of radial basis function networks) and to occur on a regular grid. This discrete approximation is a compromise which provides a high degree of flexibility yet can also be fit in practice with the available data. In many inverse problems, typical models are not feasible because they are too complex, providing more degrees of freedom

than can be fit with the scarce data available. Our approach provides enough flexibility without going too far.

We work within the Bayesian paradigm as it provides a framework for better accounting of uncertainty, particularly in the context of computer experiments and inverse problems (e.g., Kennedy and O’Hagan, 2001). As with many inverse problems, a range of initial states can produce similar fits for the data, so that the problem here is underspecified (see for example the discussions in Oliver et al., 1997; Lee et al., 2002). The Bayesian approach naturally allows us to find a range of highly plausible initial states, which can be reported through posterior distributions or intervals.

The next section provides details on the accelerator experiment, including the physical setup and the data that are collected. This section is followed by an introduction to our statistical model, some results, and finally some comments and future directions.

2 Particle Accelerator Experiments

The Low Energy Demonstration Accelerator (LEDA) is a 6.7 MeV proton accelerator that was developed to study high-current beams. It was a linear accelerator with 52 sets of magnets to focus the beam over a space of 11 meters. One of the primary areas of investigation for the physicists is the halo effect, the tendency of some of the particles to spread away from the central area of the beam. These particles can end up leaving the beam and escaping into the surroundings, creating a radioactivity hazard with repeated use of the accelerator. A better understanding of the halo forming process would allow counter measures to be taken, in particular the focusing magnets could be adjusted to reduce the halo. Such focus-

ing techniques require good understanding of initial state of the protons as they leave the accelerator, so that the magnets can be positioned properly. Thus the statistical inference for the initial distribution is a key component of the entire process.

The initial state of the beam is determined by the distribution of the particles that comprise the beam, in particular the cross-sectional distributions (in the x and y directions) of position and momentum, denoted (x, p_x, y, p_y) . While the x and y dimensions can be treated as independent, the position and momentum are expected to be correlated within a dimension, as discussed below.

If the initial distributions of position and momentum are known, the future paths of the particles can be predicted with reasonable accuracy using physical models (for example, with the Vlasov and Poisson equations) for the particle movement, as well as modeling external forces on the particles and the inter-particle Coulomb field (Dragt et al., 1988). These models are solved numerically with computer code. The scientists at Los Alamos have supplied us with the program MLI 5.0 that simulates the particle paths in such a manner (Qiang et al., 2000). A typical high fidelity run on a SunBlade 1000 workstation takes about six minutes. A lower fidelity run (with only 8,000 particles instead of 100,000) runs in about two minutes, and tends to produce similar results. Thus the simulator is reasonably fast, but is clearly the limiting factor in iterative methods.

After the beam emerges from the accelerator, it is focused with a series of magnets called *quadrupoles*. Magnets are used in sets of pairs, with the first pair focusing the beam in the y direction, but de-focusing in the x direction. The second pair focuses x but de-focues y . Through iterative focusing and de-focusing, the beam is gradually controlled in both directions. Figure 1 shows a simulation (from computer code) of the 5-th, 15-th, ..., 95-

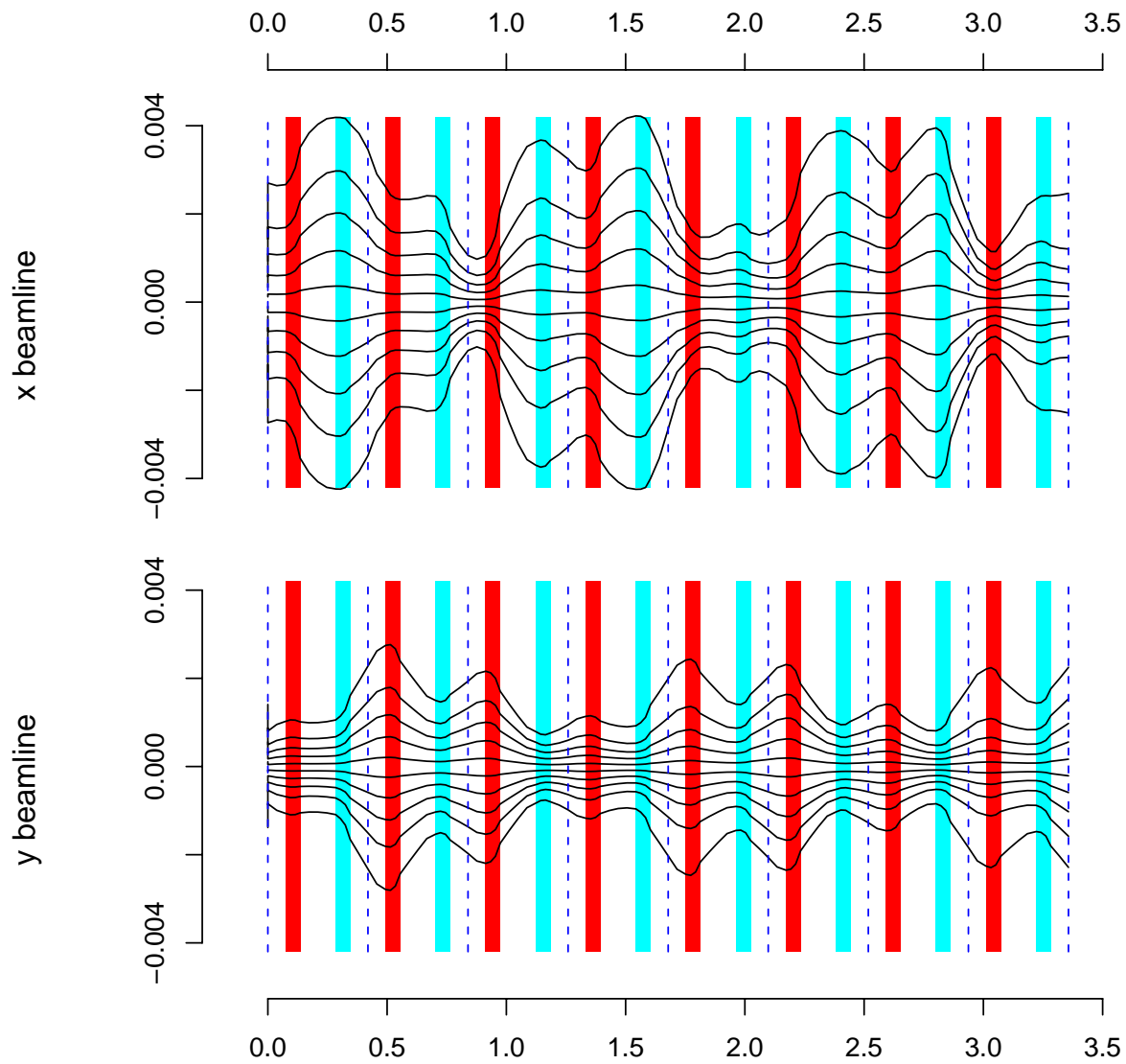


Figure 1: Simulation of particle beamlines as they are focused by a series of quadrupole magnets. The upper plot shows the progression of the particles in the x dimension, the lower plot the y dimension. Magnets are denoted by shaded areas, with alternating magnets focusing and de-focusing in each direction. The wires are shown with dashed lines.

th percentiles of the positions of particles in the beam as they pass through eight pairs of magnets. The magnets that focus in the y direction are shown in red (dark gray in grayscale), and the x focusing magnets are shown in blue (light gray). In the figure, the beam can be seen widening in the x dimension as it passes through the red magnets and narrowing as it passes through the blue ones, with the reverse effect in the y dimension.

This focusing affects the beam in that it induces a negative correlation between the position and momentum, i.e., particles further away from the center are pushed back toward the center, with the force being a function of the distance. This action is a consequence of Maxwell's equations for electromagnetic fields. In contrast, de-focusing induces a positive correlation. The net result of repeated focusing and defocusing can produce interesting non-linear effects, which provides challenges in studying the halo effect.

It is important to note that although the simulator works by tracking the progress of individual particles as they move through the apparatus, this particle-level detail can only be done via simulation. In a real experiment, particles cannot be observed individually, nor can we even jointly observe the position and momentum information. Instead, the data available are histograms of particle positions at fixed locations along the beam. The dashed vertical lines in Figure 1 represent wirescans. Physical wires are set up and as the beam crosses these wires, electrical current is produced and a histogram (here with 256 bins, which is sufficient resolution to make it appear like a curve) of particle positions can be created. These wirescans in each dimension provide the only observable data. Figure 2 provides some additional illustration. The top row are the first, third, fifth, seventh, and ninth (final) wirescans in the x direction, with the bottom row showing the corresponding y wirescans. The figure also shows the simulated particle clouds. The second row shows the x -position

versus x -momentum distribution at these five wirescan locations, and the fourth row are the corresponding distributions for the y position and momentum. The middle row shows the joint x and y positions. Notice that in this middle row the positions are independent across dimensions, whereas within dimensions the position and momentum are dependent.

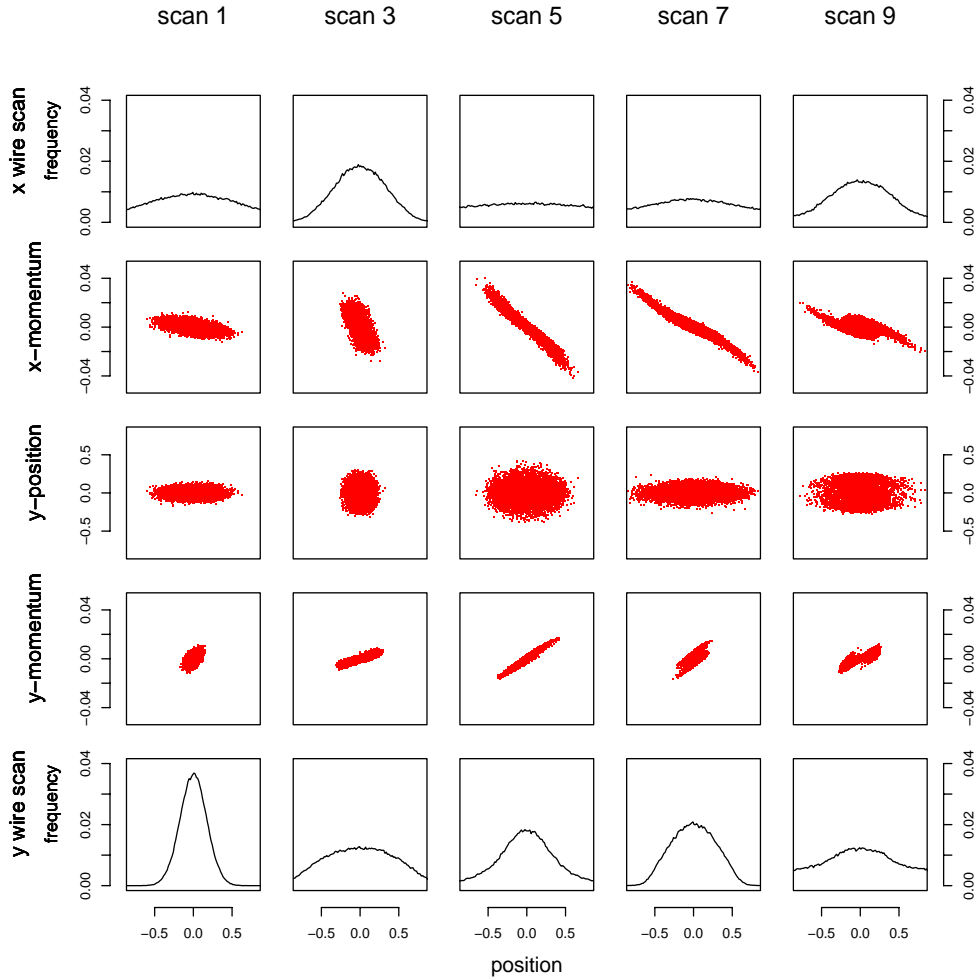


Figure 2: Simulation of wirescans and particle clouds at the locations of wirescans 1, 3, 5, 7, and 9 for the same experiment as in Figure 1. Top row is the x wirescans, second row is x -momentum vs. position, third row is y -position vs. x -position, fourth row is y -momentum vs. position, and bottom row is the y wirescans.

The plots in these two figures all fit together. For example, at the start of the beam (the left side) in Figure 1 is the first wirescan, where the trajectories are shown as somewhat spread apart for x and relatively tight for y . The upper left plot of Figure 2 shows the resulting spread out x wirescan and the lower left shows the highly peaked histogram for y position. After passing through two pairs of magnets, the second column of Figure 2 shows the wirescans and particle clouds at the third wirescan (third dashed line from the left in each panel of Figure 1). Now the x positions are relatively tight while the y distribution is wider. Note also the non-linear behavior of the particle clouds, showing intriguing relationships between position and momentum in each direction.

We want to emphasize that in practice we can only observe wirescans, and that none of the true trajectory or particle cloud information is available. In fact, the Heisenberg uncertainty principle declares that it is not possible to measure both the position and momentum of a particle. So all of our statistical inference must be based only on the wirescan position histograms (the top and bottom rows of Figure 2).

3 Statistical Model

Because the wirescans provide the only observable data, we must build our likelihood from the difference between the observed wirescans and those predicted by the computer simulator for a particular initial distribution. Thinking nonparametrically, our unknown parameter is the four-dimensional initial distribution of position and momentum in each of the x and y dimensions, i.e., the distribution of (x, p_x, y, p_y) . For a given value of this distribution, we can simulate particles from this distribution, run the particles through the simulator, and

obtain predicted wirescans $\hat{\mathbf{w}}$. These predictions are then compared to the observed data \mathbf{w} .

We use a Gaussian error structure for the difference:

$$L = \left(\frac{\tau^2}{2\pi} \right)^{2*n_{scan}*n_{bin}/2} \exp \left\{ -\frac{\tau^2}{2} \left((\mathbf{w}_x - \hat{\mathbf{w}}_x)^T (\mathbf{w}_x - \hat{\mathbf{w}}_x) + (\mathbf{w}_y - \hat{\mathbf{w}}_y)^T (\mathbf{w}_y - \hat{\mathbf{w}}_y) \right) \right\} \quad (1)$$

where \mathbf{w}_x is the wirescan data for the x dimension, \mathbf{w}_y is for the y dimension, τ^2 is a precision parameter, n_{scan} is the number of wirescans in each dimension, and n_{bin} is the number of bins per wirescan. We use an independent error structure, similar to an L^2 distance metric, as we have found that it represents a reasonable tradeoff between model details and computational efficiency. In some particular cases, there can be strong dependencies in the errors (Lee et al., 2005). In other cases, the errors show much less dependence and are effectively modeled by an independence structure. Here we have found that assuming independence produces results that are comparable to those from more complex models, so we keep the analysis more tractable by using an independent error structure.

An alternative approach to direct use of the simulator is to build a fast surrogate model (also referred to as a meta-model or a statistically equivalent model), a statistical model of the simulator which can then be run cheaply Sacks et al. (1989); Kennedy and O’Hagan (2001); Santner et al. (2003); Fang et al. (2006). In our case, this approach is not tractable because of the high dimension of the input space, which is two bivariate distributions and so theoretically of infinite dimension. Typical surrogate models are built on relatively low-dimensional input spaces, so that approach is not realistic here.

3.1 Nonparametric density estimation from a convolution perspective

Since we are essentially interested in density estimation for the initial distribution, it would seem that we have a variety of approaches to choose from in the literature. While we make no claims to have tried all of them, we have tried a number of standard approaches and found them unproductive for this application. The key difficulty is that we want to do density estimation for unobservable data. The initial distribution of states for the particles is inherently unobservable, so we must base our inference on the histograms of some of the dimensions of the transformed particle cloud. While we want enough flexibility in our model to accommodate a wide variety of possible initial distributions, we also cannot allow too much freedom or else the parameters become impractical to estimate. In essence, the amount of model flexibility provided by modern parametric and non-parametric density estimation methods results in an underspecified problem. We must also take into account that each evolution of (1) requires several minutes. Our approach balances flexibility with the ability to fit the model in a practical length of time using the data available in this problem.

In wanting a maximally flexible model, we base our approach on the convolution representation of a Gaussian process, as a Gaussian process is itself a highly flexible non-parametric model for spatially-related data in multiple dimensions. A Gaussian process is a random process $Z(\mathbf{s})$ such that for any finite subset of indices $\mathbf{s}_1, \dots, \mathbf{s}_n \in \mathbf{S}$, $Z = (Z(\mathbf{s}_1), Z(\mathbf{s}_2), \dots, Z(\mathbf{s}_n))^T$ has a multivariate normal distribution. It is common to assume a constant or polynomial form for the mean, and to assume that the covariance structure is stationary and isotropic, so that the correlation between any two points depends only on

their distance. A further common assumption is that this correlation is a simple parametric function of distance (Cressie, 1993; Wackernagel, 1998; Stein, 1999).

An alternative method of obtaining a stationary Gaussian process is by convolving white noise with a smoothing kernel (Thiébaux and Pedder, 1987; Barry and Ver Hoef, 1996; Higdon, 2002). Let $x(\mathbf{s})$ be a white noise process (or Wiener process “derivative”, e.g., Priestley, 1981), and let $k(\cdot; \phi)$ be a kernel, possibly depending on a low dimensional parameter ϕ . Convolution with this kernel produces a Gaussian process:

$$Z(\mathbf{s}) = \int_{\mathbf{S}} k(\mathbf{u} - \mathbf{s}; \phi)x(\mathbf{u})d\mathbf{u} = \int_{\mathbf{S}} k(\mathbf{u} - \mathbf{s}; \phi)d\mathbf{W}(\mathbf{u}) \quad (2)$$

where \mathbf{W} is a Wiener process. $Z(\mathbf{s})$ is such that, for $\mathbf{s}, \mathbf{s}' \in \mathbf{S}$ and $\mathbf{d} = \mathbf{s} - \mathbf{s}'$,

$$c(\mathbf{d}) = \text{cov}(Z(\mathbf{s}), Z(\mathbf{s}')) = \int_{\mathbf{S}} k(\mathbf{u} - \mathbf{s}; \phi)k(\mathbf{u} - \mathbf{s}'; \phi)d\mathbf{u} = \int_{\mathbf{S}} k(\mathbf{u} - \mathbf{d}; \phi)k(\mathbf{u}; \phi)d\mathbf{u} ,$$

i.e., $c(\mathbf{d})$ is the convolution of the kernel with itself (Kern, 2000). In the case that \mathbf{S} is \mathbb{R}^p and $k(\mathbf{s})$ is isotropic, then the resulting process $Z(\mathbf{s})$ is also isotropic, with covariance $C(\mathbf{d})$ depending only on the magnitude of \mathbf{d} . Furthermore, under standard regularity conditions on the kernel, there is a one to one relationship between the choice of kernel and the covariance function (Thiébaux and Pedder, 1987). In particular, for a given $c(\cdot)$, $k(\cdot; \phi)$ can be obtained as the inverse Fourier transform of the square root of the spectral density of c .

In practice, a discretized version of (2) is used as a close approximation. Defining the background process on a grid $\mathbf{s}_1, \dots, \mathbf{s}_m$ and suitably normalizing the kernel gives

$$Z(\mathbf{s}) \approx \sum_{i=1}^m k(\mathbf{s}_i - \mathbf{s}; \phi)x(\mathbf{s}_i), \quad (3)$$

where $x(\cdot)$ is white noise. This approximation provides a computationally efficient mechanism for using Gaussian processes in practice, particularly for larger datasets or inverse

problems, because of the reduction in effective dimension of the parameterization (Higdon et al., 2003).

Here we seek to take advantage of this lower-dimensional approximate representation of a highly flexible process. We need to modify this model because the proton distribution must be non-negative, and must be a proper density that integrates to one. We thus replace the white noise process $x(\cdot)$ with a normalized gamma process

$$x'(\mathbf{s}_i) \stackrel{iid}{\sim} \Gamma(\alpha, \beta)$$

$$x(\mathbf{s}_i) = \frac{x'(\mathbf{s}_i)}{\sum_{i=1}^m x'(\mathbf{s}_i)}.$$

The joint set of points \mathbf{x} thus has a Dirichlet distribution, but we avoid calling $x(\cdot)$ a “Dirichlet process” because that terminology already has another well-established meaning. We also note that additional flexibility can be gained by allowing the α and β parameters to vary by location. In either case, the discrete convolution gives

$$Z(\mathbf{s}) = \sum_{i=1}^m k(\mathbf{s}_i - \mathbf{s}; \phi)x(\mathbf{s}_i) . \tag{4}$$

The resulting process $Z(\cdot)$ is thus approximately a non-parametric density estimate, where the actual parameters being fit are $\{x(\mathbf{s}_1), \dots, x(\mathbf{s}_m)\}$. By fixing α , β , and ϕ , we obtain a method that retains much flexibility, but is also of low enough effective dimension that we can actually fit the model for our proton accelerator application. In choosing a Gaussian kernel, we obtain a special case of several other methods. This approach can be seen as a radial basis network with basis functions restricted to a grid, or equivalently as kernel density estimation with the kernels of identical width and restricted to a grid. The advantage of viewing our approach in the context of convolutions is that it gives us a theoretical framework to see

that much of the original non-parametric flexibility is retained provided the grid is not too coarse (Barry and Ver Hoef, 1996; Higdon, 2002).

Thus our prior model for x position and momentum is a two-dimensional process Z_x of the form (4), and our model for the y position and momentum is another such process, Z_y , independent of the x process. We use a 32 by 32 grid for each background \mathbf{s} process, which we found was the smallest grid that provided enough flexibility for a good fit; larger grids did not much improve fits but took significantly longer to run. Based on conversations with scientists at Los Alamos and on experience with simulated examples, we set $\alpha = 1$, $\beta = 50$, and the kernel k is a bivariate normal with mean zero and covariance matrices for x and y :

$$\Sigma_x = \begin{bmatrix} 3.9 * 10^{-4} & 0 \\ 0 & 4.8 * 10^{-7} \end{bmatrix} \quad \Sigma_y = \begin{bmatrix} 3.4 * 10^{-4} & 0 \\ 0 & 4.3 * 10^{-7} \end{bmatrix}.$$

This prior is then combined with the likelihood from (1) to obtain the posterior. The value of the precision, τ^2 , in the likelihood is set a priori to $1.2 * 10^{-5}$. Ideally we would fit this parameter from the data. However, as with many inverse problems (see, for example, Oliver et al., 1997; Lee et al., 2002), the data do not provide enough information to directly estimate this parameter. Thus τ^2 must be fixed, and we set it to $1.2 * 10^{-5}$. This choice represents a combination of knowledge from subject area experts and practical computing considerations, in that its value affects convergence properties of the MCMC sampler. Decreasing τ^2 forces the model to produce predictions which more closely match the observed wirescans. However, small values of τ^2 also cause the MCMC sampler to move very slowly through the parameter space. As each iteration takes several minutes, it can be impossible to obtain a good run in a reasonable amount of time (such as two weeks). Thus the choice of τ^2 is a practical compromise between expected scientific accuracy and computability.

3.2 Model Fitting

As the likelihood cannot be written in closed form, but must be evaluated numerically by running the particle simulator, we use a Metropolis-Hastings algorithm (see for example, Gamerman, 1997) to explore the posterior distribution of the background processes and hence the two position-momentum space distributions.

The standardized latent values $\mathbf{z}(\mathbf{s})$ are fit via Metropolis-Hastings by doing a random walk on the log transformed samples. For each iteration, evaluating the likelihood requires running the simulator, which requires a sample of particles. New values for the latent process $\mathbf{z}(\mathbf{s})^*$ are proposed, and conditioned on these new values, we draw a sample of size 8,000 from 1,024 lattice cell locations with the probabilities specified by the distribution induced by the convolution of the $\mathbf{z}(\mathbf{s})^*$ process for each of the x and y dimensions independently, as per Equation (3). Thus, we obtain a sample of size 8,000 of (x, p_x, y, p_y) from the product of convolutions. Next we run the proposed initial configuration of particles through the simulator to obtain a fitted wirescan. Now the likelihood (1) can be evaluated, and the new likelihood is compared to that of the previous iteration. The proposed values are accepted according to the standard Metropolis-Hasting probabilities.

As it would be rather time-consuming and computationally complex to update each latent value $\mathbf{z}(\mathbf{s})_i$ for each lattice cell individually for both the x and y dimensions, we collect latent values $\mathbf{z}(\mathbf{s})$ into several blocks according to their positions in space and update each group separately for both dimensions.

4 Simulated Example

To check the validity of our proposed approach, we examine its behavior on a simulated example, so that we can check our results against a known truth. As we have been told by scientists at Los Alamos to expect some clumping in the phase space distributions (position versus momentum), we created the true distribution to be a mixture of two normals for each of the x and y phase spaces. These distributions are shown in the bottom row of Figure 3. A cloud of particles was generated from this joint distribution and run through the simulator to produce wirescan data, and then we used only that wirescan data to try to infer the starting distribution. The top row of Figure 3 shows our posterior mean fitted distributions. We are pleased that our nonparametric methodology finds a good match when the truth is known.

Figure 4 shows additional results about the fit for this example. In that figure, each column represents a wire. The top two rows are for the x dimension, and the bottom two rows show the y dimension. The first and third rows are realizations from the estimated posterior distributions of the particle cloud (x or y position and momentum) as the beam passes through and is re-shaped by the quadrupole magnets. The second and fourth rows show the true wirescan data (circles), the estimated posterior mean scans (solid line), and posterior interval estimates (dashed lines). The posterior mean is quite close to the truth, and the credible intervals provide a measure of our uncertainty. The intervals are relatively narrow, and we conjecture this is because of a combination of two factors. First, the ground truth is a fairly “simple” distribution which can be fit relatively cleanly by our model. Second, there is some non-identifiability in the model, meaning that different input distributions can lead to very similar wirescans, so that uncertainty in the distributions may appear diminished

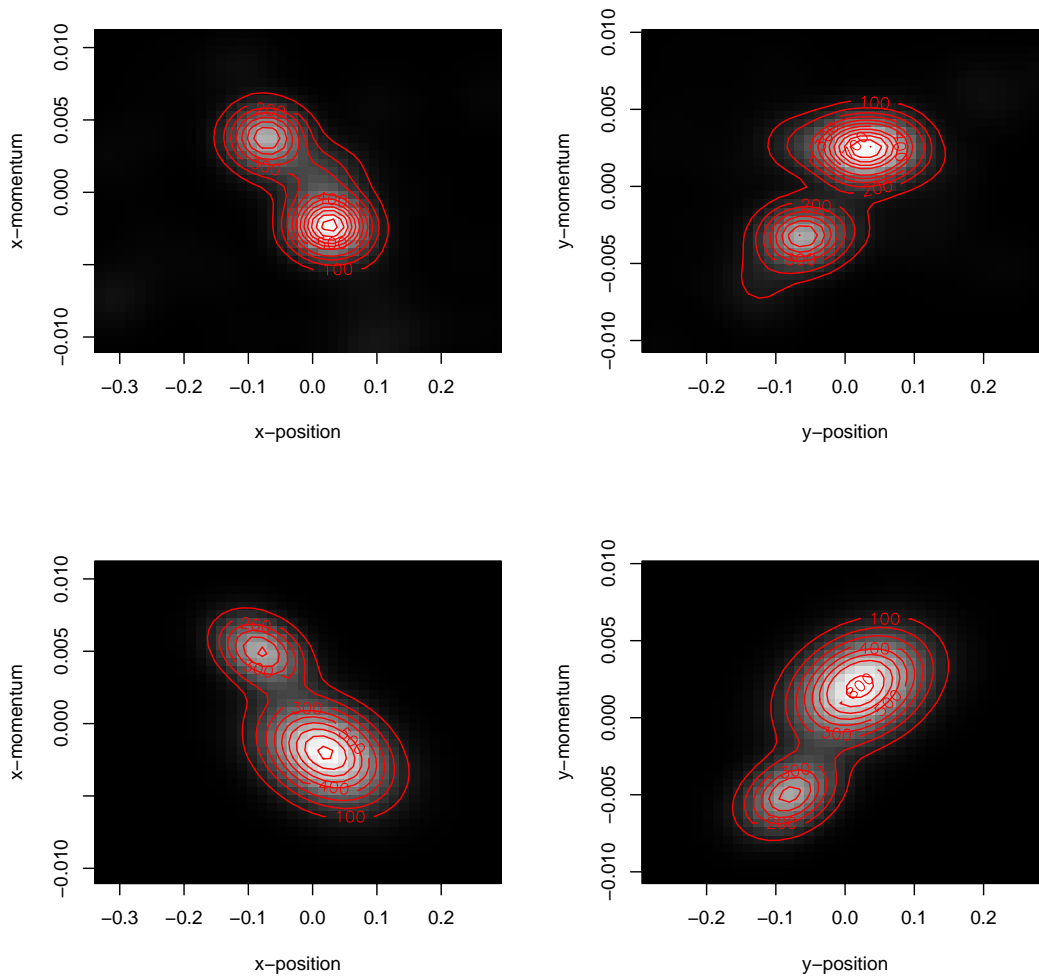


Figure 3: True and estimated input phase spaces for the simulated example.

in fitted wirescan plots.

5 Application to Real Data

Returning to the motivating application, we now present our results on data from a real run of a proton accelerator. Scientists at Los Alamos National Laboratory have provided us

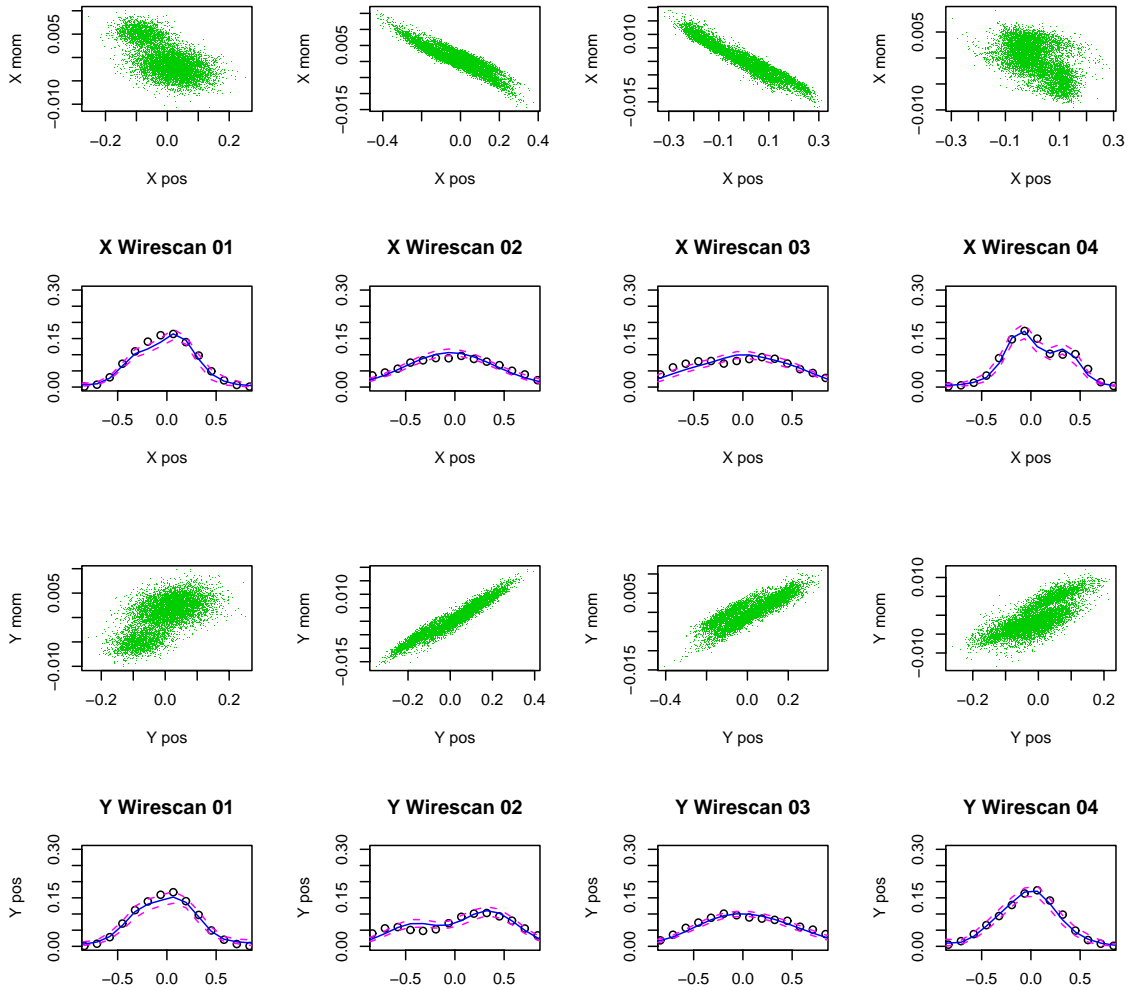


Figure 4: Results for the simulated dataset: The columns are the positions of the four wirescans. The top row shows the estimated posterior particle cloud distributions for x position vs. momentum. The second row shows the data (circle), posterior mean (solid line), and posterior 95% interval estimates (dashed lines). The third and fourth rows are the analogous plots for the y dimension. with four pairs of wirescan data (one member of each pair for the x coordinates, one for y) (Allen et al., 2002; Allen and Pattengale, 2002). From these scans we are to infer the initial

distribution of the beam in terms of its x and y position and momentum distributions.

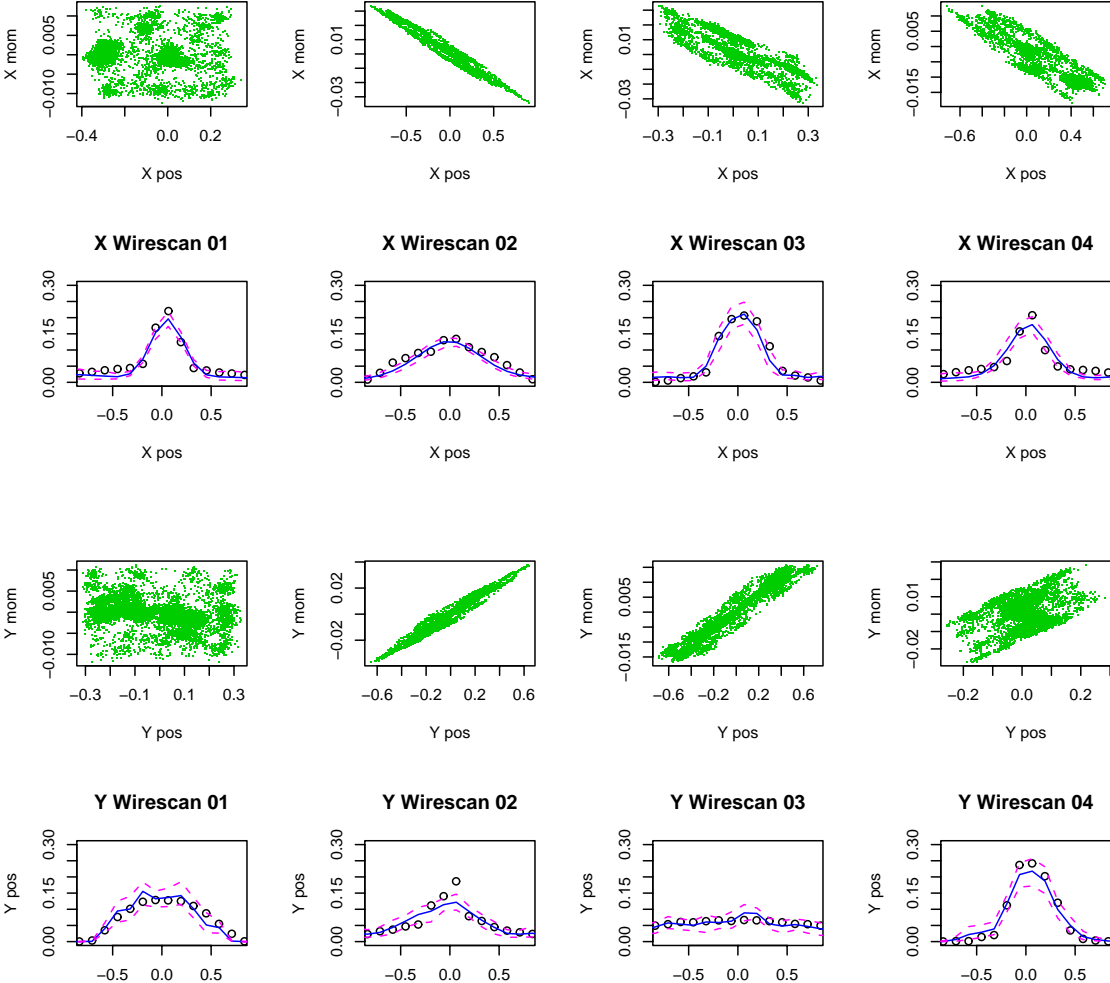


Figure 5: Results for the real dataset: The columns are the positions of the four wirescans. The top row shows the estimated posterior particle cloud distributions for x dimension vs. momentum. The second row shows the data (circle), posterior mean (solid line), and posterior 95% interval estimates (dashed lines). The third and fourth rows are the analogous plots for the y dimension.

Figure 5 shows our results on this dataset. As in Figure 4, each column is a wire, the top

two rows are for x , the bottom two for y , with estimated particle cloud posteriors and fitted scans with 95% credible intervals for each. The fitted wirescans are reasonably close to the observed values, indicating a good fit. Figure 6 shows the resulting initial distributions of position and momentum for each of the x and y dimensions. In both cases there is a primary mode representing most of the mass, with several auxiliary modes. We have been told by collaborators at Los Alamos that separate “clumps” of mass are to be expected in such plots. We also note that fitting via MCMC provides a full posterior for the initial distributions, so that uncertainty can be propagated in planning for future analysis of the beam.

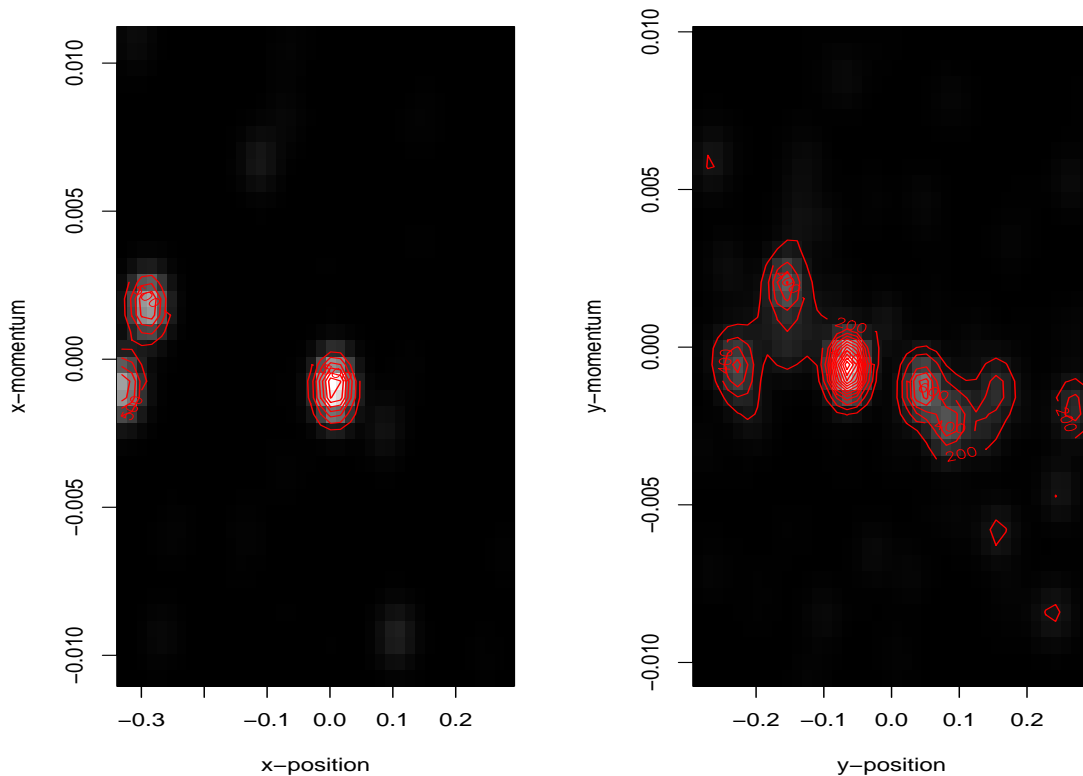


Figure 6: Estimated input phase spaces for the real dataset.

6 Conclusions

Inverse problems are often challenging because no data are directly observed on the process for which inference is to be drawn. In such cases, one must strike a balance between using an over-simplified, yet fully-identifiable model, and using a fully-flexible but under-identified model. For our proton accelerator application, we have found that traditional density estimation techniques provide more flexibility than our limited data can allow us to fit. Instead, we have presented a process convolution perspective to describe our new approach which retains much of its original nonparametric flexibility yet is still a model that can be fit from our available data.

Because the likelihood is not in closed form, uncertainty estimates are typically difficult to obtain for inverse problems. Taking the Bayesian approach provides a straightforward method of estimating our uncertainty. Such estimates are helpful for the scientists in establishing design conditions of future experiments.

Acknowledgments

The authors would like to thank Charles Nakhleh and Vidya Kumar for their helpful comments and suggestions. This work was partially supported by National Science Foundation grants DMS 0233710, DMS 0504851, and Los Alamos National Laboratory grant 74328-001-03 3D.

References

Allen, C. K., Chan, K. C. D., Colestock, P. L., Crandall, K. R., Garnett, R. W., Gilpatrick, J. D., Lysenko, W., Qiang, J., Schneider, J. D., Schulze, M. E., Sheffield, R. L., Smith,

- H. V., and Wangler, T. P. (2002). “Beam-halo measurements in high-current proton beams.” *Physical Review Letters*, 89, 214802.
- Allen, C. K. and Pattengale, N. D. (2002). “Theory and technique of beam envelope simulation: Simulation of bunched particle beams with ellipsoidal symmetry and linear space charge forces.” Tech. Rep. LA-UR-02-4979, Los Alamos National Laboratory.
- Barry, R. P. and Ver Hoef, J. M. (1996). “Blackbox Kriging: Spatial Prediction Without Specifying Variogram Models.” *Journal of Agricultural, Biological, and Environmental Statistics*, 1, 297–322.
- Cressie, N. A. C. (1993). *Statistics for Spatial Data, Revised Edition*. New York: John Wiley and Sons.
- Dragt, A., Neri, F., Rangarjan, G., Douglas, D., Healy, L., and Ryne, R. (1988). “Lie Algebraic Treatment of Linear and Nonlinear Beam Dynamics.” *Annual Review of Nuclear and Particle Science*, 38, 455–496.
- Fang, K.-T., Li, R., and Sudjianto, A. (2006). *Design and Modeling for Computer Experiments*. Boca Raton: Chapman & Hall/CRC.
- Gamerman, D. (1997). *Markov Chain Monte Carlo*. London, UK: Chapman and Hall.
- Higdon, D. (2002). “Space and Space-time Modeling Using Process Convolutions.” In *Quantitative Methods for Current Environmental Issues*, eds. C. Anderson, V. Barnett, P. C. Chatwin, and A. H. El-Shaarawi, 37–56. London: Springer-Verlag.
- Higdon, D. M., Lee, H., and Holloman, C. (2003). “Markov chain Monte Carlo-based approaches for inference in computationally intensive inverse problems.” In *Bayesian Statistics 7. Proceedings of the Seventh Valencia International Meeting*, eds. J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West. Oxford University Press.
- Kennedy, M. C. and O’Hagan, A. (2001). “Bayesian Calibration of Computer Models.” *Journal of the Royal Statistical Society, Series B*, 63, 425–464.
- Kern, J. C. (2000). “Bayesian process-convolution approaches to specifying spatial dependence structure.” Ph.D. thesis, Duke University, Durham, NC 27708.
- Lee, H., Higdon, D., Bi, Z., Ferreira, M., and West, M. (2002). “Markov Random Field Models for High-Dimensional Parameters in Simulations of Fluid Flow in Porous Media.” *Technometrics*, 44, 230–241.
- Lee, H., Sansó, B., Zhou, W., and Higdon, D. (2005). “Inferring Particle Distribution in a Proton Accelerator Experiment.” *Bayesian Analysis*. To appear.

- Oliver, D. S., Cunha, L. B., and Reynolds, A. C. (1997). “Markov Chain Monte Carlo Methods for Conditioning a Permeability Field to Pressure Data.” *Mathematical Geology*, 29, 1, 61–91.
- Priestley, M. B. (1981). *Spectral Analysis and Time Series*. San Diego, USA: Academic Press.
- Qiang, J., Ryne, R. D., Habib, S., and Decyk, V. (2000). “An Object-Oriented Parallel Particle-In-Cell Code for Beam Dynamics Simulation in Linear Accelerators.” *Journal of Computational Physics*, 163, 434–451.
- Sacks, J., Welch, W. J., Mitchell, T. J., and Wynn, H. P. (1989). “Design and Analysis of Computer Experiments.” *Statistical Science*, 4, 409–435.
- Santner, T. J., Williams, B. J., and Notz, W. I. (2003). *The Design and Analysis of Computer Experiments*. New York: Springer-Verlag.
- Stein, M. L. (1999). *Interpolation of Spatial Data: Some Theory for Kriging*. New York: Springer-Verlag.
- Thiébaux, H. J. and Pedder, M. A. (1987). *Spatial Objective Analysis with Applications in Atmospheric Science*. London: Academic Press.
- Wackernagel, H. (1998). *Multivariate Geostatistics*. Berlin: Springer.
- Yeh, W. W. (1986). “Review of Parameter Identification in Groundwater Hydrology: the Inverse Problem.” *Water Resources Research*, 22, 95–108.