Introduction

In this note we outline a procedure for determining the mass density of material surrounding the muon detector by comparing measured muon fluxes with reference or background data-sets and adjusting a set of parameters which describe the density to achieve a best overall agreement between signal and background fluxes.

Information on muon fluxes is stored in “data-cubes” which are histograms giving the numbers of detected or simulated muon tracks $n_t$ accumulated over a given period of time in bins of the measured track parameters $t = \{\phi_t, \cot \theta_t, z_t, b_t\}$\footnote{Muon tracks are described by $\phi$, the azimuth of the track pointing from the detector back toward source of the muon; $\cot \theta$, the cotangent of the zenith angle of the track; $z$, the vertical coordinate of the track from the center of the detector when it passes closest to the detector axis; and $b$, the perpendicular distance of the track from the detector axis at its point of closest approach.} [1, 2]. The distribution and density of material surrounding the detector will affect the measured values of $n_t$. Using techniques of computed tomography (CT), measured fluxes can be used to unfold material density $\rho_v$ in suitably chosen 3D voxels (indexed by $v$) located in the vicinity of the detector. To do so, one compares the numbers of tracks $n_t$ found in data-cube cells with fluxes (described by $b_t$) determined from reference or background sets where the voxel material densities are described by reference values $\rho_{0v}$. The unfolding procedure outlined here seeks to determine differences in densities $\Delta \rho_v = \rho_v - \rho_{0v}$. Various approaches are suggested for determining background fluxes needed to infer densities:

Change-detection Data-sets taken during different periods of time can be compared for changes in the surrounding medium. The set $n_t$ refers to the “signal” set; $b_t$ refers to the “background” sample. Changes in the material density will show up in the unfolded values of $\Delta \rho_v$. 

DRAFT October 7, 2008
**Flat-field** Data taken with the detector exposed to the open sky provides a reference set $b_t$ where no material is present. Scenes with objects of interest can be compared to this “flat-field” background set and the unfolded values $\Delta \rho_v = \rho_v$ give the surrounding material density relative to that of air.

**Simulation** A high-fidelity model of the detector performance, muon interactions in matter, and incident cosmic-ray muon distribution can be used to compute an estimate of the observed muon flux for a given distribution of matter surrounding the detector. Differences $\Delta \rho_v = \rho_v - \rho_{0v}$ between the actual mass density $\rho_v$ and that assumed by the simulation $\rho_{0v}$ can be unfolded from the measured muon flux $n_t$ and the simulated flux $b_t$.

The same mathematical procedure should be capable of unfolding $\Delta \rho_v$ from $n_t$ and $b_t$ in all three cases.

**$n_t$ is a Measure of Flux**

The data-sets to be compared $n_t$, $b_t$ are directly related to the muon flux measured at the detector. The conventional definition of (time- and energy-integrated) flux is the number of particles $n$ per unit phase-space element,

$$\Phi_t = \frac{d^4n}{d\Omega_t \, dA_t} = \frac{n_t}{\Delta \Omega_t \, \Delta A_t}$$

where $n_t$ is the number of muon tracks found in the phase-space volume, $\Delta \Omega_t \, \Delta A_t$, with $\Delta A_t$ oriented normal to the muon direction.

The phase-space volume is related to data-cell bin-sizes by the following Jacobians:

$$\Delta \Omega_t = \sin^3 \theta \Delta \phi \Delta \cot \theta$$  

$$\Delta A_t = \sin \theta \Delta z \Delta b$$

From these relations, it follows that the time-integrated muon flux is related to the number of entries $n_t$ in each track-parameter-bin of the data-cube by:

$$\Phi_t = n_t \left[ \frac{(1 + \cot^2 \theta)^2}{\Delta \phi \Delta \cot \theta \Delta z \Delta b} \right]$$

2
Effect of Density Differences

Muon tomography is based on the energy lost to ionization by relativistic muons in matter. When passing through ordinary materials, some of the muons headed toward the detector will "range-out", losing enough energy to fall below the threshold energy $E_{\text{th}}$ needed for detection.* Thus, fewer muons will be observed coming from regions of greater mass density.

We treat the “open-sky” as the source of muons incident on the volume of material to be imaged. Muons with energies relevant for tomography lose energy primarily by ionization, which is described by:

$$\Delta E = \int_{\text{path}} \rho(\text{path}) \left( \frac{1}{\rho} \frac{dE}{dx} \right) dl$$ (5)

where the integral is taken along the path of the muon from the open sky to the detector, $\rho(\text{path})$ is the mass density of the material along the path, and $(1/\rho dE/dx)$ is the specific ionization, which can be approximated by a constant value ($\approx 2 \text{ MeV/gm/cm}^2$) for the muon energies and materials of interest here.

The incident flux of muons at a particular zenith angle will have a spectrum of energies $E_\mu$ that follows an approximate power law form:

$$\frac{d\Phi_\mu}{dE_\mu} \sim E_\mu^{-p}$$ (6)

with a spectral index $p \approx 2.7$ [3]. To be detected, the “open-sky” muon energy $E_\mu$ must satisfy:

$$E_\mu \geq E_{\text{th}} + \Delta E$$ (7)

It is through this requirement on incident muon energy that the total amount of material along paths of muons affects the numbers of tracks detected:

$$n_{\text{det}} = \int_{E_{\text{th}} + \Delta E}^{\infty} \frac{dn}{dE_\mu} dE_\mu$$ (8)

Assume the mass density surrounding the detector in a reference or background data-set is $\rho_0$. In general, the corresponding energy losses and count-

---

*Normally, scintillator-counter or drift-tube based detectors will respond to muons with kinetic energies greater than about 0.1 GeV; adding a threshold-cerenkov detector, absorbing material, or other devices to the detector system can raise $E_{\text{th}}$ to more than 1 GeV. For reference, a relativistic muon will lose about 0.5 GeV per meter of rock.
ing rates are:

\[ E_{\text{ref},t} = \int_{t} \rho_0 \left( \frac{1}{\rho} \frac{dE}{dx} \right) dl \]  

(9)

\[ b_t = \int_{E_{\text{th}} + E_{\text{ref},t}}^{\infty} \frac{dn}{dE_\mu} dE_\mu \]  

(10)

where the path labeled by \( t \) corresponds to an average over the phase-space corresponding to cell \( t \) of the data-cube.

Next consider the change in \( b_t \) that occurs when one voxel (labeled by \( v \)) differs in density by \( \Delta \rho_v \) from the reference value. To first order, the difference in the number of counts \( \Delta n_{tv} \) is given by:

\[ \Delta n_{tv} = -l_{tv} \Delta \rho_v \left( \frac{1}{\rho} \frac{dE}{dx} \right) \left. \frac{dn}{dE_\mu} \right|_{E_{\text{th}} + E_{\text{ref},t}} \]  

(11)

where \( l_{tv} \) is the path-length of track \( t \) through voxel \( v \).

Under the assumption that the muon energy spectrum is described by a power law, the value of the differential energy spectrum evaluated at the minimum detectable energy limit is simply related to the total number of detected tracks, the threshold energy, and the spectral index \( p \). Therefore, the difference expected in the number of detected counts \( \Delta n_t \) in phase-space element \( t \), relative to the reference number of counts \( b_t \), is given by:

\[ \frac{\Delta n_t}{b_t} = -\sum_v l_{tv} \Delta \rho_v \left( \frac{1}{\rho} \frac{dE}{dx} \right) \left( p - \frac{1}{E_{\text{th}} + E_{\text{ref},t}} \right) \]  

(12)

where \( l_{tv} \) now represents the average path-length through voxel \( v \) encountered by an ensemble of tracks spanning cell \( t \) of the data-cube. Eqn. 12 gives the connection between voxel densities and data-cube flux differences that will be used to unfold the density of material surrounding the detector.

To simplify expressions that follow, we choose to combine terms in Eqn. 12 into the weights \( l_{tv} \) and a “voxel density amplitude” \( k_v \) given by:

\[ k_v = \Delta \rho_v \left( \frac{1}{\rho} \frac{dE}{dx} \right) \left( p - \frac{1}{E_{\text{th}} + E_{\text{ref},t}} \right) \]  

(13)

These density amplitudes can be used as fitting parameters to unfold the 3D structure of the material surrounding the detector. The reference energy-loss \( E_{\text{ref},t} \) depends on the reference density \( \rho_0 \) and path \( t \); in some cases, we
expect $E_{\text{ref},t}$ to vary smoothly over the phase-space region of interest. In such cases, the density amplitudes alone are probably sufficient for producing useful 3D imaging. In cases, where the $t$ variation is significant, Eqn. 12 is still appropriate and the density differences $\Delta \rho_v$ can be used directly as fitting parameters.

Another way to express $k_v$, when the reference density is significantly greater than that of air, shows explicitly that it is a measure of the contrast in counting rate associated with the contrast in mass density provided by the voxel:

$$k_v = \frac{\Delta \rho_v}{\int_t \rho_0 \, dl} \left( \frac{p - 1}{1 + E_{\text{th}}/E_{\text{ref},t}} \right)$$

Adding $\Delta n_t$ to $b_t$, the expectation value $\lambda_t$ for number of counts $n_t$ in cell $t$ based on the reference background set $b_t$ is related to the voxel density amplitudes $k_v$ and path-length weights $l_{tv}$ by:

$$\lambda_t = r_{sb} b_t \left( 1 - \sum_v l_{tv} k_v \right)$$

where $r_{sb}$ is an overall normalization factor that gives the relative running times or exposures for the signal ($n_t$) and background ($b_t$) data-sets. $r_{sb}$ could be varied as an additional fitting parameter when comparing the two data-sets to determine the values of $k_v$.

The next section outlines a procedure to compare data ($n_t$) with estimates ($\lambda_t$) to derive density information. Because $\lambda_t$ is derived directly from $b_t$, it is important that the two data sets $n_t$ and $b_t$ be recorded under experimental conditions as nearly identical as possible. When the $b_t$ are taken with the same detector, either through flat-fielding or under change-detection, it is relatively easy to achieve operating conditions that are similar. One must ensure that PMT gains, electronic threshold settings, and other such experimental details are the same for both sets of runs. In this way, triggering efficiency, track reconstruction efficiency and the actual geometry which determines the phase-space volume of data-cube cells should be the same for both data-sets. When the set $b_t$ is derived from a simulation, it is essential that accurate models of detector performance and the cosmic ray energy spectrum and angular distribution be employed.

### Likelihood Fit

As noted in the introduction, the strategy for unfolding voxel density differences $\Delta \rho_v$ or amplitudes $k_v$ is to find values for these parameters (and the
normalization ratio \( r_{sb} \) that lead to the best joint description of the data-
sets \( n_t \) and \( b_t \). We propose to do this by maximizing a likelihood function
\( L \) based on the statistical agreement between \( n_t \) and \( \lambda_t \) for all cells under
consideration.

The numbers of tracks found in each cell of the data-cube are described
by Poisson statistics with a probability distribution \( p_n(\lambda) = \lambda^n \exp(-\lambda)/n! \)
that \( n \) counts are observed when the average number expected is \( \lambda \). Our
proposed likelihood function is given by the product of the individual Poisson
probabilities for all data-cells being compared:

\[
L = \prod_t p_{n_t}(\lambda_t)
\]  \hspace{1cm} (16)

Rather than dealing directly with \( L \), the fitting parameters can be opti-
mized by minimizing the negative-log-likelihood function \( \mathcal{L} = -2 \ln L \). \( \mathcal{L} \)
plays a role here similar to a least-squares function for fitting data that have
Gaussian errors. Because the \( n_t! \) terms that enter \( L \) do not depend on any
fitting parameters, \( \mathcal{L} \) can be further simplified to: \( \dagger \)

\[
\mathcal{L} = 2 \sum_t (\lambda_t - n_t \ln \lambda_t)
\]  \hspace{1cm} (17)

where the \( \lambda_t \) are functions of \( b_t \) and the fitting parameters \( \{k_v\} \) and \( r_{sb} \) (see
Eqn. 15).

This choice of likelihood function is based on the assumption that the
estimators \( \lambda_t \) are well behaved and do not introduce additional statistical
fluctuations. Inspection of Eqn. 15 shows that negative values of \( \lambda_t \), while
unphysical, are possible depending on values picked for \( k_v \). Therefore, care
must be taken when minimizing \( \mathcal{L} \) to ensure physically meaningful values for
the estimators by imposing the constraint \( \lambda_t > 0 \). Furthermore, the \( \lambda_t \) will
normally have statistical errors because they are derived from detected or
simulated background data-sets \( b_t \), also having Poisson statistics. If possible,
the exposure or simulation producing \( b_t \) should be much larger than the
exposure taken for the signal event sample \( n_t \); equivalently, it is desirable
to be able to compare data-sets where \( r_{sb} \ll 1 \). In any event, care must
be taken with zeros that may be present in the \( b_t \) set used to determine
\( \lambda_t \) to avoid numerical problems while optimizing \( \mathcal{L} \). The simplest thing

\( \dagger \)It was noted above that the relative normalization parameter \( r_{sb} \) could be fitted along
with \( \{k_v\} \) to minimize \( \mathcal{L} \). Such a fit is equivalent to imposing the constraint \( \sum_t \lambda_t = \sum_t n_t \). This can also be done using the method of Lagrange multipliers by adding a term
\( \mu \sum_t (\lambda_t - n_t) \) to Eqn. 17 with the additional fitting parameter \( \mu \).
to consider would be to only include cells where \( b_t \geq N_{\text{min}} \), say 10 counts minimum.

Treating the density differences \( \Delta \rho_v \) as a perturbation, the first approximation \( \{ k^{(1)}_v \} \) to the solutions of Eqn. 17, using the \textit{ansatz} given in Eqn. 15, can be determined by solving the \textit{linear} equations:

\[
\sum_u \left[ \sum_t l_{tv} n_t l_{tu} \right] k^{(1)}_u = \sum_t l_{tv} (r_{sb} b_t - n_t)
\]

which have the same form as an ordinary least-squares problem. As expected, the covariance matrix \( \left[ \sum_t l_{tv} n_t l_{tu} \right]^{-1} \) provides estimated statistical uncertainties for the solutions \( \{ k^{(1)}_v \} \) based on \( \sqrt{n_t} \).

\section*{Some Details}

The computational effort needed to carry out the program outlined here will depend on details of the detectors and voxels to be unfolded, such as the following:

- The currently operating “large” prototype detector stores data into data-cubes having a total of 15.6M bins, 600 in \( \phi \), 130 in \( \cot \theta \), 20 in \( z \), and 10 in \( b \). It is expected that the 1/3-scale detectors, now under construction, will use data-cubes with 1/9 the number of cells: the angular resolution will be approximately the same as the large prototype, while the projected area is 1/9-th that of the prototype. Such an arrangement has the same phase-space volume per data-cube cell for either type of detector.

- The method outlined here works with multiple detectors. Each detector can be thought of as simply adding more data-cube cells and corresponding muon counts to the data-sets to be unfolded.

- We can expect that a typical data run will involve the equivalent of running the large prototype detector for about 10 weeks which will accumulate roughly 1-2 G tracks or about 100 counts per average data-cube cell.

- A \((0.5 \text{ m})^3\) voxel located 10 m radially and 10 m above the detector will “illuminate” approximately 2400 data-cube cells.

- Regarding computational load, it should be noted that the path-length weights \( l_{tv} \) need be computed only once for a particular choice of voxels and date-cube cells to be unfolded.
Conclusions

A plausible scenario for extracting 3D mass density information from detected muon tracks exists. Now, it needs to be confirmed with simulations and real data!

References

