Post-processing modeling and removal of background noise in space-based time-of-flight sensors

Daniel. J. Gershman,[1] Jason A. Gilbert,[1] Jim M. Raines,[1] George Gloeckler,[1] Patrick Tracy,[1] and Thomas H. Zurbuchen[1]

Department of Atmospheric, Oceanic, and Space Sciences, University of Michigan, Ann Arbor, MI, USA

This paper develops and implements a mathematical framework that enables noise modeling and removal for any sensor that relies on discretely measured events. Here, we apply this technique to data from the Fast Imaging Plasma Spectrometer (FIPS), a time-of-flight mass spectrometer on the MESSENGER spacecraft. An iterative Monte Carlo event-processing algorithm is used to probabilistically separate instrument measurements into real and noise-based events. Kernel density estimation is employed as a smoothing technique to enable noise removal for datasets comprised of only a few events. Given an accurate noise model, the overall misidentification of events is expected to be less than 25% even for datasets having low signal-to-noise (SNR) ratios, with substantially improved results expected for progressively larger accumulations of data. These techniques are shown to successfully recover heavy ion events from in-flight FIPS data both inside and outside Mercury's magnetosphere. Such data analysis methods not only drive a more in-depth understanding of sensor operation, but also provide a unique post-processing approach that can result in the improvement of in-flight SNR without any modifications to instrument settings. In addition, these method are readily applicable to existing archived datasets of past missions.

I. INTRODUCTION

Noise events in a spaceborne time-of-flight mass spectrometer (TOF-MS) are undesirable measurements that can arise from a variety of sources, ranging from background processes to incident particle events. These erroneous events contaminate mass and energy spectra, making it difficult to recover accurate ion composition and abundances, especially for heavy ion species for which measured event rates can be low. A number of these noise sources, as well as their effects on data analysis are introduced and discussed in detail by Gilbert et al.[1]. Here, we focus on modeling and probabilistic removal of these events, such that the signal-to-noise ratio (SNR) of measured datasets can be increased using software post-processing techniques rather than adjusting limited in-flight instrument settings. The techniques described here were used to enable a number of scientific analyses[2–5] of observations from the Fast Imaging Plasma Spectrometer (FIPS)[6,7] on the MErcury Surface, Space ENvironment, GEochemistry, and Ranging (MESSENGER) spacecraft[8], currently in orbit around Mercury.

There are two general classes of noise contamination: additive and multiplicative/distortion. Additive noise is an unwanted signature superimposed on top of a real signal[9]. The most common and perhaps ubiquitous form of this contamination is white noise[10], with a constant spectral density and a Gaussian distribution of amplitudes typically superimposed on an audio, electronic, or image signal. For any additive noise source, however, the underlying signal remains unchanged. Distortion, on the other hand, involves a convolution of the uncontaminated signal with a noise function, altering its original properties[11]. A simple example of distortion is image blurring, whereby the color in a particular pixel is changed on the basis of the original properties of its neighbors[12].

The vast majority of noise events in a TOF-MS are additive. Although, as will be discussed, the intensity of some noise sources will scale with the number of real events, these noise events are nonetheless superimposed on the measured distribution of incident particles. Therefore, TOF-MS noise removal typically requires only distribution separation rather than dataset deconvolution. Also, the TOF-MS noise sources discussed here will be characterized with accumulations of instrument data and will consequently be well-known. A known source is considerably easier to remove from a dataset than an unknown one, and consequently we will not require the use of complex blind-signal separation techniques[13]. Finally, the ability to post-process measured datasets enables the devotion of substantial computational resources

to the removal process, allowing each event to be analyzed individually and completely, as opposed to being limited by the real-time processing power of an on-board flight computer.

In the case of many audio and image processing applications, noise sources exhibit specific spectral signatures[14]. After a frequency analysis of a particular dataset, high- or low-pass filtering can be used to reduce markedly the contribution of many noise sources[15]. This strategy, however, may not be best suited for applications to TOF-MS measurements that are a set of discrete, multi-dimensional events for which properties must not be altered as a consequence of any noise reduction process. Instead, we will employ Monte Carlo techniques[16] to process instrument-measured events. Monte Carlo techniques have been applied to a wide variety of applications in fields such as statistical physics[17], biology[18], and finance[19]. Although these techniques vary in their implementation, they all use a set of sampled random numbers to aid in their computations. Here, the probability that a particular event is noise will be derived with an instrument-noise forward model. These probabilities, in conjunction with randomly generated numbers will be used to process sets of TOF-MS observations.

The primary obstacle for accurate noise reduction for TOF-MS data is low measured event rates. TOF-MS measurements are discrete datasets that often result in zero recorded events for a given time step, as opposed to an analog audio signal or a digital image. In fact, most measured TOF-MS data will rarely form a smooth distribution function unless they are accumulated over long time periods. The challenge here is to reconstruct a smooth distribution function sufficient for event processing based on only a few discrete events.

Data from MESSENGER/FIPS are used as examples for both noise forward modeling and removal processes. FIPS is a double-coincidence TOF-MS, composed of an electrostatic analyzer (ESA) and a TOF section. As illustrated in Figure 1, charged particles are guided by an electric field between shaped electrodes of the ESA such that only ions with a particular energy per charge (E/q) ratio successfully reach and penetrate a thin carbon foil, liberating secondary electrons. These electrons are accelerated by electric fields in the TOF section and electrostatically reflected by a mirror harp assembly onto a microchannel plate (MCP) detector, where they create a start signal. The incident particle, which is most often neutralized by the foil[20], passes through the mirror harp assembly to another MCP detector, generating a stop signal, resulting in a correlated TOF event. The anode in the start MCP assembly is position-sensitive, such that valid events have five properties: time step t, energy
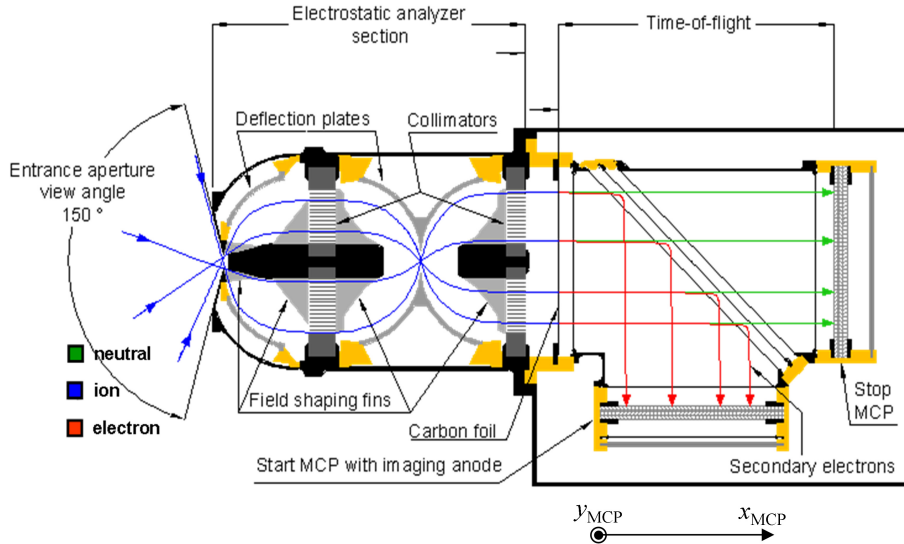
FIG. 1. Illustration of the MESSENGER/FIPS TOF-MS adapted from Andrews et al. [7]. Ions with a particular range of E/q are filtered through an ESA and then analyzed by a TOF section with imaging capabilities on the start detector. The y_{MCP} coordinate direction of the MCP is out of the plane of the figure.

per charge E/q, TOF, and start MCP coordinates $(x_{\mathrm{MCP}}, y_{\mathrm{MCP}})$, with coordinate system geometry defined in Figure 1.

Although FIPS is used as a case study, the techniques developed here should be readily generalizable to any TOF-MS. In Section II, we develop a mathematical framework for TOF-MS measurements as well as a noise removal process suitable for datasets with varying SNR and low total count rates. In Section III, we discuss a method to create a complete TOF-MS noise forward model from accumulations of measured data, using MESSENGER/FIPS as an example. Finally, in Section IV, we apply the noise forward modeling and removal processes to successfully recover measured heavy ion events in and around the magnetosphere of Mercury from MESSENGER/FIPS in-flight observations.

II.   MATHEMATICAL FRAMEWORK

In order to model noise in a TOF-MS sensor, a mathematical framework for describing how an instrument measures events must be developed. Such a framework will enable comprehensive noise forward modeling and removal processes. The techniques described here

TABLE I. Notation used for the mathematical framework of TOF-MS noise modeling and removal processes. Here, Φ is a dummy variable used to represent any of several different quantities and distributions introduced throughout the paper.

Notation	Definition
Φ_{real}	Real signal distribution of Φ
Φ_{noise}	Noise signal distribution of Φ
Φ_{bg}	Background signal distribution of Φ
Φ_{inc}	Incident particle signal distribution of Φ
$\hat{\Phi}$	Normalized distribution of Φ
Φ_{m}	Φ derived from instrument measurement
Φ^*	Φ derived from noise-removal processing

are applicable to any TOF-MS. Table I introduces notation that is used to describe quantities throughout the paper. Additional subscripts and superscripts that appear throughout the paper but are not given in Table I will be defined at first use. The derivations here, as well as the noise processing developed later in this section, are limited to one dimension. However, as will be discussed, these results will apply to multi-dimensional datasets.

A.   Measured events

Let the distribution of all measured events with respect to some parameter, x, be defined as $f_{\mathrm{m}}(x)$. In general, x is one of the measurable event parameters (e.g., t, E/q, TOF). Since it is produced from discretely measured events, $f_{\mathrm{m}}(x)$ necessarily can only have integer values. The measured events that form $f_{\mathrm{m}}(x)$ can be considered to be a set of random samples of some probability density distribution, $f(x)$. $f(x)$ can have non-integer values, but its total integrated value must be the same as for $f_{\mathrm{m}}(x)$ such that, for all possible values of x,

$$\sum_{\mathrm{i}} f_{\mathrm{m}}(x_{\mathrm{i}}) = \sum_{\mathrm{i}} f(x_{\mathrm{i}}). \tag{1}$$

The instrument-measured $f_{\mathrm{m}}(x)$ is formed through a histogram of discretely measured events. In such an accumulation, each event corresponds to a Kronecker delta function (δ_{ij})
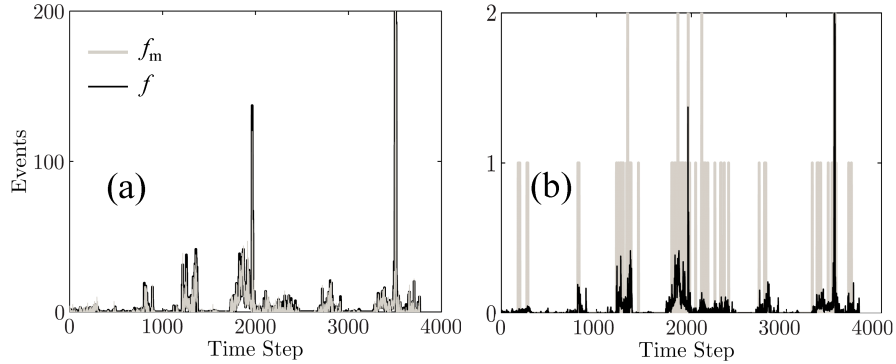
FIG. 2. Comparison of measured distribution $f_\mathrm{m}(x)$ obtained from generating N samples from the example distribution $f(x)$ for (a) $N = 10{,}000$ events and (b) $N = 100$ events. For small numbers of events there are large point-wise differences between $f_\mathrm{m}(x)$ and $f(x)$. For large N, $f_\mathrm{m}(x) \approx f(x)$. For each pair of distributions, $N = \sum_\mathrm{i} f_\mathrm{m}(x_\mathrm{i}) = \sum_\mathrm{i} f(x_\mathrm{i})$.

shifted to the appropriate histogram bin center, x_i. $f_\mathrm{m}(x)$ can be written in terms of a sum of measured events as,

$$f_\mathrm{m}(x) = \sum_{\mathrm{i}=1}^{N} \delta_{\mathrm{x,x_i}}, \tag{2}$$

where N is the total number of measured events and x_i is the histogram bin location of the ith event.

For large numbers of events, $f_\mathrm{m}(x)$ and $f(x)$ are nearly identical, as indicated in Figure 2a, which shows a sampling of a distribution $f(x)$ for $N = 10{,}000$ events. However, since $f_\mathrm{m}(x)$ must necessarily have integer values, for small numbers of events, $f_\mathrm{m}(x)$ can differ substantially from $f(x)$. This difference is illustrated in Figure 2b, which shows a sampling of only $N = 100$ events from the same distribution. Therefore, $f(x)$ can be thought of as the limit of $f_\mathrm{m}(x)$ at large N.

Distributions of events can be separated into two subsets: noise-based events and real particle events, such that,

$$f(x) = f_\mathrm{real}(x) + f_\mathrm{noise}(x), \tag{3}$$

and

$$f_\mathrm{m}(x) = f_\mathrm{m,real}(x) + f_\mathrm{m,noise}(x). \tag{4}$$

6

The total number of noise-based events is defined as η, and the total number of real events is defined as S, with $S + \eta = N$.

As will be discussed in Section III, a noise forward model will naturally produce the $f_{\text{noise}}(x)$ distribution, since average noise event rates, which can be non-integers, are used to generate the expected noise for a given time interval.

1. *Signal-to-noise ratio estimation*

In order to best characterize and understand the impact of noise sources on data analysis, a direct estimation of the instrument SNR is needed. For a given accumulation of data, the measured SNR_{m} is defined as the total number of measured real events, S, divided by the total number of measured noise-based events, η:

$$SNR_{\text{m}} \equiv \frac{\sum\limits_{\text{i}} f_{\text{m,real}}(x_{\text{i}})}{\sum\limits_{\text{i}} f_{\text{m,noise}}(x_{\text{i}})} = \frac{S}{\eta} = \frac{N - \eta}{\eta} \tag{5}$$

As discussed previously, large accumulations of data are needed to have $f_{\text{m,noise}}(x) \approx f_{\text{noise}}(x)$. However, an accumulation over one or more dataset dimensions should be sufficient to achieve $\sum_{\text{i}} f_{\text{m,noise}}(x_{\text{i}}) \approx \sum_{\text{i}} f_{\text{noise}}(x_{\text{i}})$. In the case that this relationship holds, SNR_{m} is directly calculable and can be used to obtain a direct estimate of the signal-to-noise ratio for any data product from the total measured data, f_{m}, and forward modeled noise, f_{noise}. This value can be used as a quality indicator for a set of measured data, independent of any noise removal process that may be applied.

2. *Noise event probabilities*

Following equation (3), the probability that a measured event with property x is a noise-based event can be written as,

$$P_{\text{m,noise}}(x) = \frac{f_{\text{m,noise}}(x)}{f_{\text{m}}(x)}. \tag{6}$$

Likewise, equation (4) can be used to form the probability,

$$P_{\text{noise}}(x) = \frac{f_{\text{noise}}(x)}{f(x)}, \qquad (7)$$

which is effectively the large-N limit of equation (6).

If there is a known probability, $P_{\text{m,noise}}(x)$, that a measured event with property x is noise, then a Monte Carlo processing technique can be used to separate a distribution into real and noise-based subsets as follows: (1) For each measured event, a random number is uniformly generated between 0 and 1; (2) given the property x of each event, the corresponding probability $P_{\text{m,noise}}(x)$ is compared with the value of the random number; (3) if the generated number is smaller than $P_{\text{m,noise}}(x)$, the event is flagged as noise-based. Otherwise, the event is considered to be real; (4) After all events have been processed, recovered distributions $f^*_{\text{m,noise}}(x)$ and $f^*_{\text{m,real}}(x)$ are produced. It follows from these steps that an accurate estimate of $P_{\text{m,noise}}(x)$ is vital for a successful noise removal process.

Consider several sets of N measured events randomly sampled from the distribution $f_{\text{noise}}(x)$. Although each set originates from the same distribution, for small N the point-wise differences between their sampled distributions will be large. No set of events is necessarily better than another; they are simply independent samplings of the same distribution. It is therefore extremely difficult to precisely reproduce an instrument-measured $f_{\text{m,noise}}(x)$ distribution from random samplings of $f_{\text{noise}}(x)$, as all possible combinations of samplings are computationally unreasonable to calculate. Therefore, it is not feasible to directly determine $P_{\text{m,noise}}(x)$.

P_{noise}, however, can be readily computed because f_{noise} is directly produced from a noise forward modeling process (Section III), and $f(x)$ can be accurately estimated for a given $f_{\text{m(x)}}$ (Section II B). Therefore, as an approximation, P_{noise} rather than $P_{\text{m,noise}}$ is used in order to create a computationally reasonable Monte Carlo processing algorithm. To take into account errors in the estimation of P_{noise}, this Monte Carlo technique can be iterative, whereby the dataset is continually reprocessed using updated values of $f(x)$ and f_{noise} until the total expected amount of noise has been removed, or until the algorithm converges. As detailed in Appendix A, with the use of synthetic datasets to characterize algorithm accuracy, this process will necessarily remove the amount of noise predicted by f_{noise} but will result in the misidentification of some events (less than 20% of all events even for low SNR values). The real events that are more likely to be misclassified, however, are those that are statistically indistinguishable from the distribution of noise events. Concentrations

of high-SNR events, for example, are less likely to be misclassified than single events during times of low SNR.

B.   Kernel density estimation

In order to determine $P_{\text{noise}}(x)$, the distributions $f(x)$ and $f_{\text{noise}}(x)$ must be known. Noise forward modeling will produce $f_{\text{noise}}(x)$, and for large N, $f_{\text{m}}(x)$ naturally approaches $f(x)$. However, for small N, another technique must be used to convert the measured $f_{\text{m(x)}}$ to a reasonable approximation of $f(x)$. Kernel density estimation (KDE)[21] is a technique that can accomplish this conversion for small N. In KDE, instead of using delta functions as in equation (2), each event is represented by an arbitrary normalized function, or kernel. Kernels that span multiple histogram bins enable non-integer representation of measured events. A Gaussian kernel is a natural choice for TOF-MS systems since measured events often accumulate to form a set of Gaussian-like peaks due to their normally distributed error in the high-count limit.

A KDE-based distribution of measured events, $f_{\text{m,KDE}}(x)$, can be written as a sum of normalized discrete Gaussians:

$$f_{\text{m,KDE}}(x) = \frac{\sum\limits_{\text{i}=1}^{N} e^{-\left(\frac{x-x_{\text{i}}}{w(x)}\right)^2}}{\sum\limits_{\text{x}=-\infty}^{\infty} \sum\limits_{\text{i}=1}^{N} e^{-\left(\frac{x-x_{\text{i}}}{w(x)}\right)^2}} N. \tag{8}$$

Here, w is the bandwidth of the Gaussian kernel for which the value must be tuned to a particular type of dataset. In general, w can be a function of x. Proper selection of bandwidth $w(x)$ can be difficult and is distribution dependent[22]. This choice is further complicated by the fact that the distribution of TOF-MS data changes with time. For this reason, $f_{\text{noise}}(x)$ is used to aid bandwidth selection, since the measured distribution must encompass the noise distribution. The choice of $w(x)$ here is not found analytically, but rather empirically, to provide a reasonable estimation of $f(x)$ guided by measured instrument data.

Our empirical bandwidth will be composed of two parts. First, we recognize that integer values of expected noise, on average, should be accompanied by a corresponding integer number of measured events. The fractional part of $f_{\text{noise}}(x)$, however, will likely be accom-

9

panied with either zero or one measured events. For example, if the expected number of noise events is 3.3, we would expect to measure either 3 or 4 noise events. Therefore, the kernel bandwidth should be based on the fractional part of $f_{\text{noise}}(x)$, $f_{\text{noise}}(x) - \lfloor f_{\text{noise}}(x) \rfloor$, where $\lfloor\ \rfloor$ denotes the integer part of the enclosed quantity[23]. This factor is well suited for setting bandwidths for datasets composed almost entirely of noise, i.e., $SNR_{\text{m}} \approx 0$. In this case, following equation (8), the amplitude of the Gaussian kernel for a single event will be approximately equal to its corresponding noise distribution value.

For distributions with high SNR, however, a bandwidth set by this fractional noise factor will tend to overspread, resulting in a poor estimate of $f(x)$. To compensate for this over-spreading, the factor $1/(1+SNR_{\text{m}})$ is included. This factor will reduce the overall spreading of the measured events for increased signal strength, where $SNR_{\text{m}} \approx 0$ and $SNR_{\text{m}} \to \infty$ result in spreading reduction factors of 1 and 0, respectively. Together, the fractional noise and spreading factors combine to form the empirical bandwidth,

$$w(x) = \frac{1}{f_{\text{noise}}(x) - \lfloor f_{\text{noise}}(x) \rfloor} \cdot \frac{1}{1 + SNR_{\text{m}}}. \tag{9}$$

Equation (9) is intended to be generally applicable, suitable for datasets with both strongly localized and uniformly distributed events and a variety of SNR values. Further modifications to this bandwidth may be required to best fit a particular dataset of interest.

A metric that determines how well a generated distribution matches the true distribution $f(x)$ is the relative error between them. The relative error, E, is defined as,

$$E(f, f_{\text{m}}) \equiv \frac{\sum\limits_{i} |f(x_{\text{i}}) - f_{\text{m}}(x_{\text{i}})|}{\sum\limits_{i} f(x_{\text{i}})}, \tag{10}$$

where $f(x)$ is the true distribution and $f_{\text{m}}(x)$ is the measured distribution. Here, the sum over i represents only the portion of the dataset over which there are measured events, since those are the only steps where P_{noise} will be computed.

Some individual steps (x_{i}) will have a very high local SNR, for example, when a real event is measured at a step with almost no expected noise. In this case, the bandwidth $w(x_{\text{i}})$ from equation (9) will be large, causing overspreading such that $f_{\text{m,KDE}}(x_{\text{i}})$ will be very different from $f(x_{\text{i}})$. Whereas the factor $1/(1 + SNR_{\text{m}})$ can help to mitigate this overspreading, high local SNR is still possible for steps with overall SNR_{m} close to zero. However, in
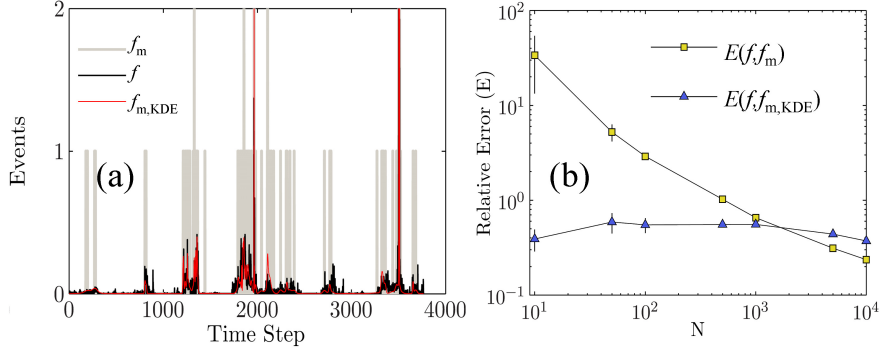
FIG. 3. (a) An example distribution $f(x)$ with corresponding f_m and $f_{m,KDE}$ distributions for $N = 100$ and $SNR_m = 0$. (b) $E(f, f_m)$ and $E(f, f_{m,KDE})$ as functions of N, showing that KDE techniques provide better estimates of f for small N, with a relative error nearly independent of N. Ten sets of samples were obtained from f for each data point. The mean and standard deviation of E for the 10 sets are shown for each as the data marker and error bars, respectively.

such cases the ratio $f_{noise}(x_i)/f_{m,KDE}(x_i)$ can still be very small, leading to $P_{noise}(x_i) \approx 0$ and the measured event not being removed, as desired. The most sensitive SNR regime for noise removal will therefore be $SNR_m \to 0$, and for such a regime it is imperative that $f_{m,KDE}(x_i) \approx f(x_i)$.

The distribution $f(x)$ from Figure 2 and its corresponding $f_{m,KDE}$ distribution with $SNR_m = 0$ are shown in Figure 3. $f_{m,KDE}$ clearly provides a better estimation of $f(x)$ than the samples forming $f_m(x)$. To quantify this improvement, $E(f, f_m)$ and $E(f, f_{m,KDE})$ are computed for various N values as shown in Figure 3b. For large N, the KDE distributions have a slightly higher relative error than f_m, but for small N they are more accurate by two orders of magnitude.

In the case of noise modeling for TOF-MS, KDE techniques can be applied for any accumulation, such as time history, TOF, or incident particle direction datasets. For the incident particle direction datasets, the bandwidth can apply to a two-dimensional Gaussian kernel. In general, KDE techniques can be applied to higher-dimensioned datasets[22], with each dimension permitted its own characteristic bandwidth. The visualization of such techniques becomes difficult, and in the case of TOF-MS, the determination of appropriate bandwidths may become intractable. For this reason, KDE techniques for TOF-MS may be practically limited to lower-dimensional datasets. As detailed in Appendix B, use of KDE techniques

11

with a Monte Carlo event processing algorithm should result in the misidentification of less than 25% of all events, only a modest increase in overall error when compared with using a known $f(x)$ instead of $f_{\mathrm{m,KDE}}(x)$ (determined as <20% in Appendix A) .

III. FORWARD MODELING OF INSTRUMENT NOISE

A good estimate of $f_{\mathrm{noise}}(x)$ is vital for an accurate noise removal. Furthermore, even in the absence of a suitable approximation for $f(x)$, $f_{\mathrm{noise}}(x)$ will provide a direction estimation of instrument SNR. The complete $f_{\mathrm{noise}}(x)$ distribution should have contributions from any and all noise sources in an instrument. Here, the MESSENGER/FIPS sensor is used as a baseline, and both background and active noise sources will be discussed and modeled appropriately to demonstrate how to generate $f_{\mathrm{noise}}(x)$ from a set of accumulated observations.

An unfiltered 50-day accumulation from the day-of-year (DOY) range 100–149 in 2011 of all MESSENGER/FIPS-measured events is shown in Figure 4. The E/q of each event is plotted as a function of TOF, with curved tracks corresponding to real events that are superimposed onto various noise sources. This time period of data will serve as the calibration dataset that will be used to characterize the behavior of f_{noise}. Data from the E/q steps marked as "Background" will be used to characterize the background noise rates and distributions. The distributions of events from the $E/q - TOF$ steps marked as "Protons" will be used to model the contamination of the H^+ mass peak into the peaks for species of higher mass per charge (m/q) ratios. Finally, the events within the $E/q - TOF$ steps marked as "ESD" will provide the rates of events triggered by increased electron-stimulated desorption (ESD) of material from the start MCP at times of high incident particle flux. The distribution of each of the active noise sources will be extrapolated to the TOF region 200–300ns, a range over which neither source appears to be dominant.

A. Passive noise sources

Background, or passive noise, is constant in time and independent of E/q. However, such noise may have specific TOF and MCP position dependence. Consider as an example, electron-stimulated desorption processes in the MESSENGER/FIPS sensor. Two
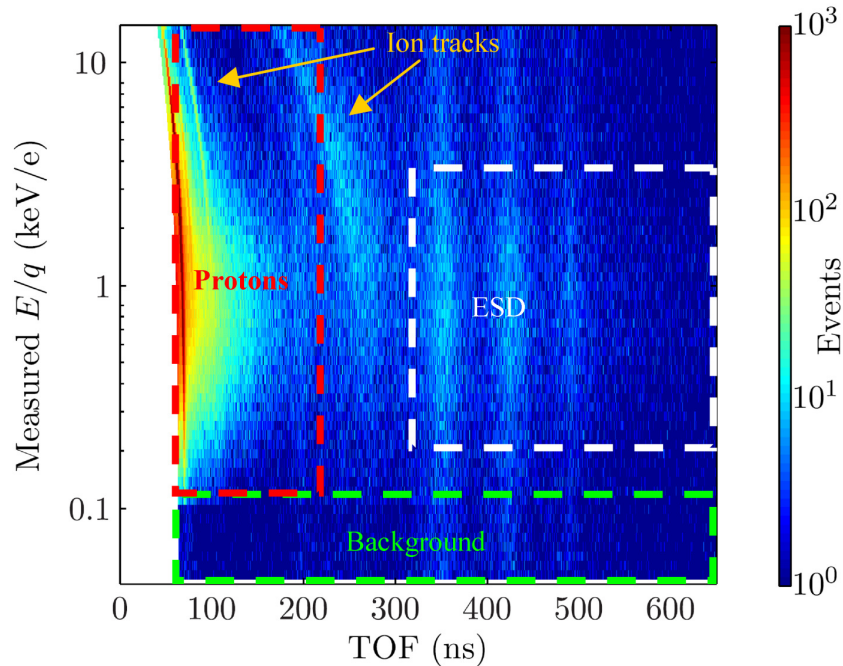
FIG. 4. (a) $E/q - TOF$ spectrogram of 50 days (day of year 100-149 in 2011) of orbital MES-SENGER/FIPS heavy-ion ($m/q > 1$ amu/e) measurements. Curved tracks correspond to incident ions. Boxes marking the $E/q - TOF$ steps corresponding to background events, proton tail events, and induced electron-stimulated desorption events are indicated as "Background", " Protons", and "ESD", respectively, and will be used to characterize each noise source.

background electron sources have been identified, as described by Gilbert et al.[1]: emission from high-voltage wire harps, and a localized field emission source near the edge of the MCP active area. A complete noise model must include distributions corresponding to each source.

The E/q steps marked as "Background" in Figure 4 are the lowest instrument energy per charge steps over which few numbers of real ion events are expected, particularly outside of Mercury's magnetosphere. To further avoid possible contamination from any actual low energy per charge events, only time steps where the ratio between the start MCP rate and valid event rate was greater than 10 are used, as real incident events will have ratios of ~2 due to the efficiencies of ion detectors. Due to frequent drifts in instrument operating conditions, most notably ambient temperature, the background rate calculations were performed from daily accumulations of data. The background rates are the "front" rate, i.e., the background

13

signal on the start MCP (FSR_{bg}); the "rear" rate, i.e., the background signal on the stop MCP (RSR_{bg}); the background proton event rate, i.e., background events with low TOF, (PE_{bg}); and the valid event, i.e., the double-coincidence, background rate (VE_{bg}). These rates (in units of s^{-1}) are used in various stages of the instrument noise forward model. The accumulation time in seconds for each time step, $\Delta(t)$, is readily calculable and can change with the operating mode of the instrument. The background TOF and MCP $x - y$ spectra are obtained using the full 50-day accumulated dataset, since they are not expected to change quickly with time.

The total background distribution must be separated into its harp emission and other field emission subcomponents, denoted with subscripts "he" and "fe" respectively. A simple separation isolates events corresponding to the "hotspot" pixels (Gilbert et al.[1]) in the background MCP $x - y$ distribution as the generalized field emission source and attributes the rest to wire harp emission of electrons. Following Gilbert et al.[1], the TOF distribution of the harp emission source corresponds to the flight times of desorbed ions from the surface of the start MCP to the top of the TOF chamber. The TOF of the field emission source also includes a set of peaks at a constant fraction of the flight times of the harp emission source, indicating that desorbed species are striking the sides of the TOF chamber instead of the top. These accumulated distributions are smoothed to reduce statistical noise. The resulting distributions of MCP $x - y$ and smoothed-TOF for each source are shown in Figures 5 and 6, respectively.

The TOF background noise distributions $f_{\mathrm{noise,fe}}(t, TOF, x_{\mathrm{MCP}}, y_{\mathrm{MCP}})$ and $f_{\mathrm{noise,he}}(t, TOF, x_{\mathrm{MCP}}, y_{\mathrm{MCP}})$ can be written for a given accumulation of time following equations (11) and (12),

$$f_{\mathrm{noise,fe}}(t, TOF, x_{\mathrm{MCP}}, y_{\mathrm{MCP}}) = g_{\mathrm{fe}} \cdot VE_{\mathrm{bg}} \cdot \hat{f}_{\mathrm{fe}}(TOF) \cdot \hat{f}_{\mathrm{fe}}(x_{\mathrm{MCP}}, y_{\mathrm{MCP}}) \cdot \Delta(t), \qquad (11)$$

and

$$f_{\mathrm{noise,he}}(t, TOF, x_{\mathrm{MCP}}, y_{\mathrm{MCP}}) = g_{\mathrm{he}} \cdot VE_{\mathrm{bg}} \cdot \hat{f}_{\mathrm{he}}(TOF) \cdot \hat{f}_{\mathrm{he}}(x_{\mathrm{MCP}}, y_{\mathrm{MCP}}) \cdot \Delta(t). \qquad (12)$$

Here, g_{fe} and g_{he} are the fractions of contributions of both sources, where $g_{\mathrm{fe}} + g_{\mathrm{he}} = 1$. VE_{bg} is the total background valid event rate in s^{-1}, and $\Delta(t)$ is the accumulation time for a given time step in s. Other background sources of noise can be characterized similarly.
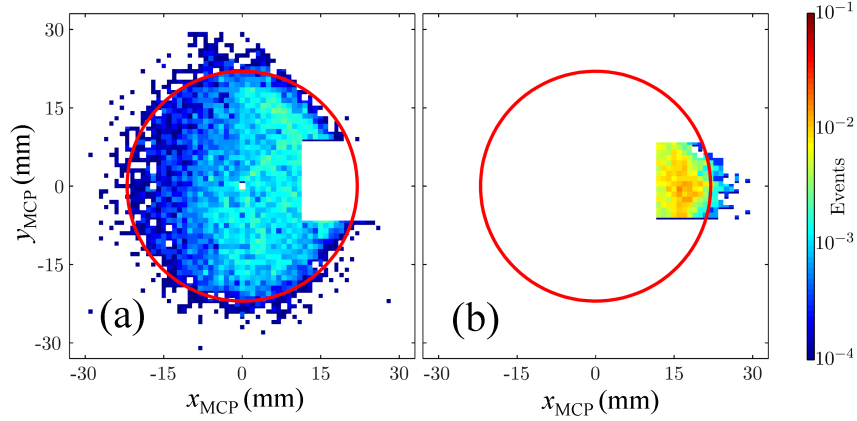
FIG. 5. Normalized MCP $x - y$ distribution of background ESD events in MESSENGER/FIPS triggered by (a) harp emission electrons ($\hat{f}_{\mathrm{he}}(x_{\mathrm{MCP}}, y_{\mathrm{MCP}})$) and (b) field emission electrons ($\hat{f}_{\mathrm{fe}}(x_{\mathrm{MCP}}, y_{\mathrm{MCP}})$). The red circle indicates the boundaries of the MCP active area. The field emission electrons are concentrated in a beam forming a set of "hotspot" pixels near the edge of the active area whereas the harp emission electrons strike almost uniformly over the surface of the MCP.
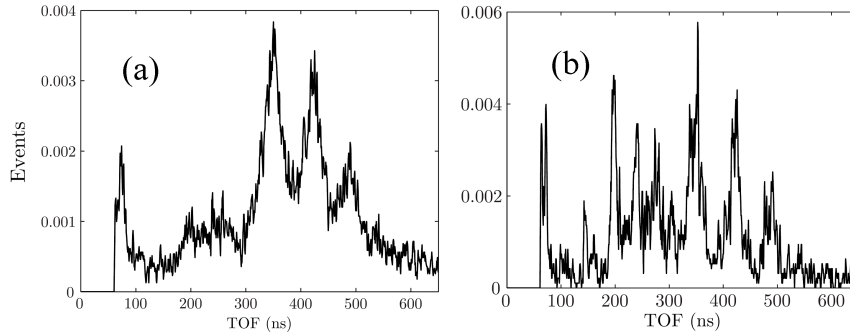
FIG. 6. Normalized TOF distribution for background ESD events in MESSENGER/FIPS triggered by (a) harp electrons ($\hat{f}_{\mathrm{he}}(TOF)$) and (b) field emission electrons ($\hat{f}_{\mathrm{fe}}(TOF)$). The harp emission TOF distribution corresponds to the flight times of desorbed species from the surface of the start MCP to the top of the TOF chamber. The field emission events include a contribution from desorbed ions that strike the side of the TOF chamber, leading to flight times at a constant fraction of the corresponding harp emission distribution peaks.

15

B.   Active noise sources

Active sources create noise as a function of incident particles or photons. Any external stimulus capable of generating signals from the start and stop detectors is considered an active source. The sources considered here will be ion energy straggling and induced ESD, as discussed in detail by Gilbert et al. [1]. Accidental coincidences due to high flux or penetrating radiation will not be modeled due to the extremely low SNR of measurements during those time periods.

1.   *Ion energy straggling*

Particles passing through a thin carbon foil lose some of their kinetic energy, a phenomenon called "energy straggling"[24]. The resulting distribution features a high TOF peak tail for each species. For a TOF-MS, H^+ will always have the smallest TOF, as it has the smallest possible m/q of 1 amu/e. The tail for H^+, therefore, has the potential to contaminate the TOF channels corresponding to all other ion species, as shown in Figure 4. Furthermore, for MESSENGER/FIPS, the most easily measured ion is H^+, which is present in both the solar wind and Mercury's magnetospheric environments and is typically orders of magnitude more abundant than other species[25]. The rates of heavier ion events can be comparable to the magnitude of the proton peak tail, making it difficult to distinguish between the two sets of particles. Events in the $E/q - TOF$ steps in Figure 4 labeled as "Protons" are assumed to be dominated by the proton peak tail.

Although it is possible to obtain an overall good fit to energy-straggled TOFs using a kappa distribution[24], for heavy ion species small errors in the fit can create relatively large errors in f_{noise}. Therefore, rather than use a functional fit for the entire proton peak shape, the accumulation of data from Figure 4 is smoothed and used as the noise distribution directly, analogous to the formation of the background noise distributions. The average proton peak shape is expected to change with time, as changes in ambient temperature may affect the energy straggling process through the carbon foil, and the thickness of the carbon foil may not be spatially uniform, leading to different peak shapes for time periods where, on average, particles enter the instrument aperture from different directions. For MESSENGER/FIPS, the proton peak shape is updated from 10-day accumulations, providing a balance between
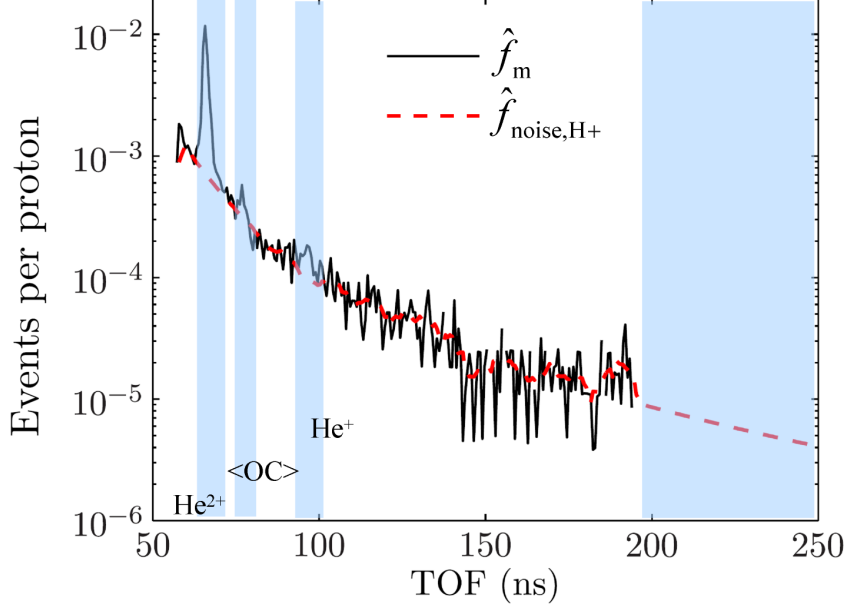
FIG. 7. TOF distribution of proton tail events from $E/q = 1.5$ keV/e for DOY 100–209 in 2011, normalized by the proton event rate. TOFs corresponding to solar wind ion tracks for He^{2+}, He^{+}, and $\langle OC \rangle$, the averaged solar wind heavy ions[3], are indicated. The black solid curve is the set of normalized raw measurements . The red dashed curve represents a smoothed curve that has been interpolated through ion track TOFs and extrapolated to higher TOFs with a power law.

statistical significance and capturing this time variation.

The expected background spectum calculated in Section III A is subtracted from this accumulation, and TOF channels corresponding to actual ion tracks are excluded. The data are linearly interpolated through these data gaps and extrapolated to higher TOF channels with a power-law fit, consistent with a kappa TOF peak shape. An example TOF spectra for a 10-day accumulation at $E/q = 1.5$ keV/e is shown in Figure 7, with a corresponding noise curve fit that interpolates through ion track locations and extrapolates to higher TOF values.

The amplitude of the H^{+} tail should scale with the total number of measured proton events (PE) measured at lower TOFs, i.e., near the peak of the H^{+} TOF distribution. The accumulated data in each energy step is therefore normalized by the total corresponding proton event rate to form $\hat{f}_{noise,H+}(TOF, E/q)$. The errors in this quantity are greatest

for energy steps in which the proton event rate is comparable to the background rate. However, since the expected noise from proton events in these steps will be low, this error is manageable. The MCP $x - y$ distribution of this noise will be that of the incident protons, $\hat{f}_{\text{inc,H+}}(E/q)$, which can be obtained by accumulating events close to the proton peak at each energy step. The expected f_{noise} therefore becomes:

$$
\begin{aligned}
f_{\text{noise,H+}}(t, TOF, E/q, x_{\text{MCP}}, y_{\text{MCP}}) = (PE(t, E/q) - PE_{\text{bg}}) \\
\times \hat{f}_{\text{noise,H+}}(TOF, E/q) \\
\times \hat{f}_{\text{inc,H+}}(x_{\text{MCP}}, y_{\text{MCP}}) \\
\times \Delta(t). \qquad (13)
\end{aligned}
$$

2. Induced ESD

In addition to electrons from field emission or harp emission sources, secondary electrons from incident particles can desorb ions from instrument surfaces, resulting in valid TOF events. For MESSENGER/FIPS, a number of ion species are desorbed from the start MCP and follow trajectories that lead to correlated stop signals. This effect occurs with some probability relative to an incident particle being detected. Therefore, the total number of these events should be proportional to the number of non-background start signals detected in a particular E/q step. Also, desorption rates will depend not only on the number of incident particles but on where those particles enter the instrument.

Given, as in Section III A, that a uniform distribution of harp electrons strikes the MCP, the harp electron event distribution is proportional to the number of valid events per start signal in each MCP "pixel". Not every start signal will have a corresponding stop signal, so the ratio of the valid event rate to the start rate will always be less than one. In addition, the incident particle distribution ($\hat{f}_{\text{inc}}$) should correspond to the distribution of start signals on the MCP. Multiplying $\hat{f}_{\text{inc}}(x_{\text{MCP}}, y_{\text{MCP}})$ by the background harp electron event distribution $\hat{f}_{\text{he,bg}}(x_{\text{MCP}}, y_{\text{MCP}})$ should therefore give an estimation of the MCP spatial distribution of induced ESD events.

The TOF distribution for induced ESD events for MESSENGER/FIPS should follow that of the harp-emission-based events. However, it should be noted that incident particles

on average tend to hit closer to the center of the MCP active area than the background events. Ions desorbed from these locations have a larger distance to travel due to the curved top of the TOF chamber. This effect is observed as a small shift in TOF (TOF_{o}) of the harp electron TOF distribution. For MESSENGER/FIPS, TOF_{o} can be computed for each daily accumulation of data through the calculation of the change in the first moment of the largest TOF peak (325–375ns) from the expected background distribution, following equation (14),

$$TOF_{o} = \frac{\sum_{\mathrm{ESD}} TOF \cdot (f_{\mathrm{m}} - f_{\mathrm{he,bg}})}{\sum_{\mathrm{ESD}} (f_{\mathrm{m}} - f_{\mathrm{he,bg}})} - \frac{\sum_{\mathrm{ESD}} TOF \cdot f_{\mathrm{he,bg}}}{\sum_{\mathrm{ESD}} f_{\mathrm{he,bg}}}, \tag{14}$$

where

$$\sum_{\mathrm{ESD}} \equiv \sum_{TOF=325\mathrm{ns}}^{375ns} \sum_{(E/q)_{\mathrm{ESD}}} \sum_{t} \sum_{(x_{\mathrm{MCP}},y_{\mathrm{MCP}})}. \tag{15}$$

The distribution of ESD-induced noise with respect to E/q step and MCP $x - y$ can then be written as,

$$f_{\mathrm{noise,ESD}}(t, x_{\mathrm{MCP}}, y_{\mathrm{MCP}}, E/q) = c \cdot (FSR(t, E/q) - FSR_{\mathrm{bg}})$$
$$\times \hat{f}_{\mathrm{he}}(t, x_{\mathrm{MCP}}, y_{\mathrm{MCP}})$$
$$\times \hat{f}_{\mathrm{inc}}(t, x_{\mathrm{MCP}}, y_{\mathrm{MCP}}, E/q)$$
$$\times \hat{f}_{\mathrm{he,bg}}(TOF - TOF_{o})$$
$$\times \Delta(t), \tag{16}$$

where c is a constant.

The constant c can be found from a daily accumulation of data such that,

$$\sum_{\mathrm{ESD}} (f_{\mathrm{m}} - f_{\mathrm{he,bg}}) = \sum_{\mathrm{ESD}} f_{\mathrm{noise,ESD}}. \tag{17}$$

With a full TOF spectrum of noise known, the ESD spectrum readily applies to the "overlap" region in TOF between 200 ns and 300 ns.
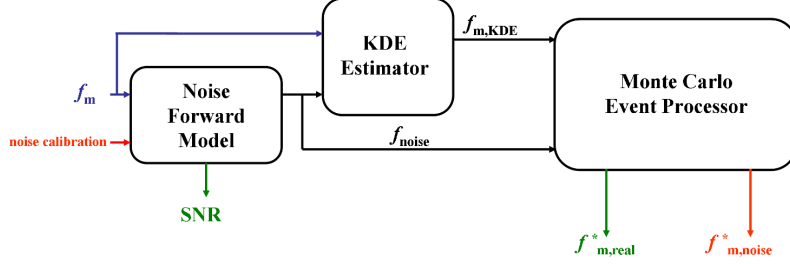
FIG. 8. Block diagram describing noise modeling and event processing for MESSENGER/FIPS. Measured data as well as noise calibration factors are input into a noise forward model that produces f_{noise} and consequently, estimates of dataset SNR. The measured data f_{m} along with f_{noise} is input to a KDE estimator that provides $f_{\text{m,KDE}}$, a distribution suitable for use in a Monte Carlo processor that separates f_{m} into $f^*_{\text{m,real}}$ and $f^*_{\text{m,noise}}$.

IV. NOISE PROCESSING FOR MESSENGER/FIPS

As illustrated in Figure 8, with all noise sources well characterized, f_{noise} can be accurately determined for a given set of measured data f_{m}, resulting in a direct estimation of SNR for any size dataset. A KDE estimator, given f_{noise} and f_{m} as inputs, can produce $f_{\text{m,KDE}}$, an approximation to the underlying distribution function (f) that can be analyzed using a Monte Carlo event-processing algorithm, and the measured distribution of real events, $f^*_{\text{m,real}}$, can be recovered. This model is applied here to in-flight data from MESSENGER/FIPS.

A. Noise forward modeling

To model the various sources of noise, the relevant noise distributions from Section III must be added together for each time step. It is important to note that the spatial distribution of incident particles is required in both the models of proton peak tail noise and induced ESD events for MESSENGER/FIPS. When the SNR in each E/q step is low, an iterative noise removal/modeling process may be required to obtain the proper incident spatial distribution. However, in the case of MESSENGER/FIPS, the number of expected noise events will be small for a single time step, i.e., $f_{\text{m,noise}}(t) \ll f_{\text{m,real}}(t)$. E/q steps that contribute substantially to active noise sources will have large numbers of incident particles that should dominate the measured MCP $x-y$ distribution. Therefore the total MCP $x-y$ distribution

20

is used for an approximation to the spatial distribution of incident particles. The error in this approximation will be large for small numbers of incident particles, but since active noise sources scale with the measured start rate, the contribution from these low count-rate steps to the total accumulated active noise should be negligible.

Due to the number of data dimensions and their corresponding interdependencies, a complete five-dimensional $f_{\mathrm{noise}}(t, E/q, TOF, x_{\mathrm{MCP}}, y_{\mathrm{MCP}})$ distribution may be difficult to manage and store. Furthermore, as discussed in Section II B, noise removal using KDE techniques is challenging for higher-dimensional datasets. Therefore, instead of producing the complete distribution of noise, several noise data products can be produced that represent accumulations along one or more of the dimensions. Such summation seemingly removes vital information for reconstructing an accurate $f_{\mathrm{m,real}}$. However, as will be discussed in Section IV B 1, with proper choice of accumulations, a reasonable $f_{\mathrm{m,real}}$ can be recovered. For MESSENGER/FIPS, the MCP $x - y$ distribution is measured with the least precision. Furthermore, the majority of the noise sources are either widely spread over the MCP active area or have a similar positional dependence as the incident particles. Therefore, ignoring the MCP $x - y$ dependence is expected to improve statistical significance of accumulations with minimal impact on the accuracy of modeled distributions.

Two accumulations are calculated here as an example. The first, f_{noise}, is a $E/q - TOF$ spectrogram for a particular time period. Such an accumulation, although not as useful from the perspective of scientific analysis, provides a good evaluation of noise forward-modeling performance. The second, f_{noise}, is an energy-dependent time series for a given ion that accumulates over MCP $x - y$ and several TOF values, and that series will be the baseline for all FIPS heavy ion science analysis.

An accumulated $E/q - TOF$ distribution (9a) with corresponding forward modeled noise (9b) for 10 days (DOYr 100–109, 2011) in shown in Figure 9. Ion tracks for He^{2+}, He^+, Na^+, and Ca^+ are marked. For clarity, not all ion tracks present in the data have been included. The two distributions vary only at the locations of ion tracks, where the noise distribution underestimates the total number of events, as expected. There are good matches with the proton peak tail and induced ESD noise in terms of both $E/q - TOF$ location and intensity. Synthetic data are used in Appendix A to demonstrate that with an accurate f_{noise}, removal should be accurate to within 25%. Given the quality of the generated f_{noise} shown in Figure 9, this estimation of error should be directly applicable to measured MESSENGER/FIPS
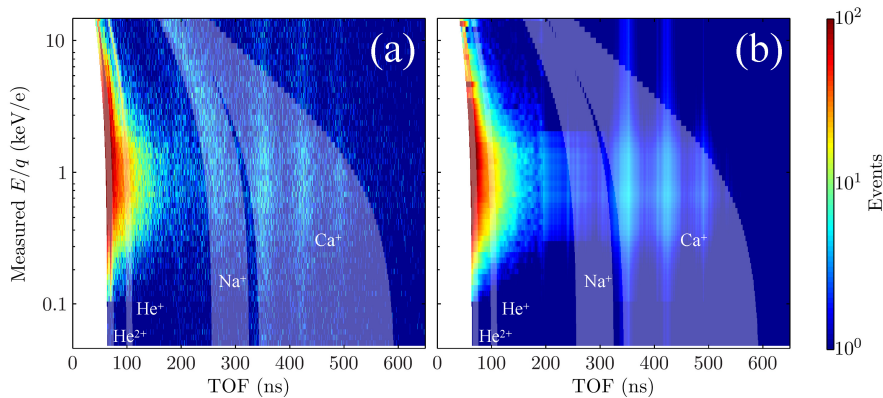
FIG. 9. (a) 10 day (DOY 100–109, 2011) $E/q - TOF$ accumulation of events measured by MESSENGER/FIPS. Tracks corresponding to heavy ions He^{2+}, He^{+}, Na^{+}, and Ca^{+} are highlighted. (b) Forward modeled $E/q - TOF$ noise distribution for the same time period. There is an excellent match between the measured data and generated f_{noise} distribution, with the only major discrepancies occurring at the locations of known ion tracks, at which actual events are expected to be present.

data.

B.   Noise removal

Once f_{noise} is generated, the set of processed events can be split into sets of recovered real and noise-based events. Although it is not possible to calculate the accuracy of this separation for measured datasets, the expected properties of recovered events, both real and noise, can be used to gain confidence in the removal process. On-orbit MESSENGER/FIPS data (DOY 100–109, 2011) have been analyzed using daily accumulations of He^{2+}, He^{+}, Na^{+}, and Ca^{+} at each E/q step, yielding a series of one-dimensional time series suitable for processing with the algorithm described in Section II A 2.

Each time series is further split into two parts: "active" and "quiet", with active times defined as periods when the proton rate goes above 15 events per E/q step (i.e., inside Mercury's magnetosheath or during FIPS solar wind observations), and quiet times (typically inside Mercury's magnetosphere) defined as when the proton rate has remains below 15 events per E/q step for at least 64 FIPS energy per charge scans (1 hour or 1 minute
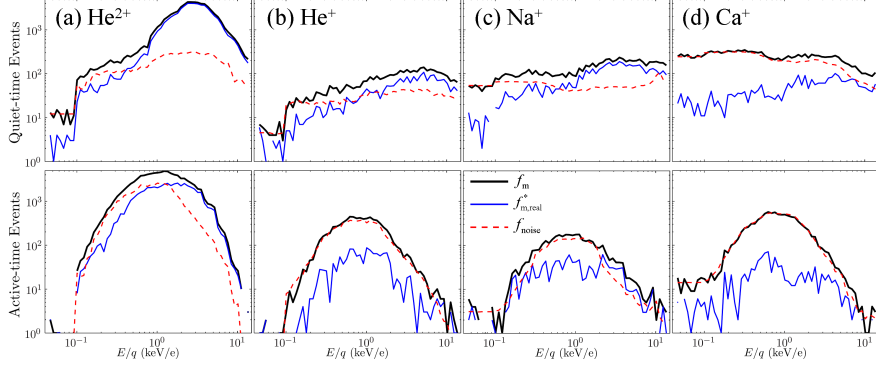
22

FIG. 10. Total measured (f_{m}), modeled noise (f_{noise}), and recovered ion events ($f^*_{\mathrm{m,real}}$) as a function of E/q for a 10 day (DOY 100–109, 2011) accumulation of MESSENGER/FIPS on-orbit data for quiet and active periods for (a) He^{2+}, (b) He^+, (c) Na^+, and (d) Ca^+. The noise processing algorithm accurately recovers ion events, resulting in a substantial increase in measurement signal-to-noise ratios, especially for E/q steps corresponding to a high incident proton flux.

depending on operating mode). The active time periods will be dominated by active noise sources such as H^+ energy straggling and induced ESD, which have larger uncertainty in the forward modeled f_{noise}. If the noise for a particular step is overestimated, real events that occur during quiet times would be thrown out. By splitting up the time series, a more sensitive recovery of events is expected.

The total number of events from both active and quiet time periods for each ion as a function of E/q is shown in Figure 10, along with corresponding recovered real events and predicted f_{noise}. For all ions, the large peak near $E/q = 1$ keV/e is effectively removed, since it corresponds to the proton peak tail/ESD events. From Figure 9, where there are clear ion tracks at higher energies in the absence of modeled noise, almost no events are removed.

Statistics will also play a role in the recovery, as the counting error ($\sqrt{N}$) for low SNR datasets results in some uncertainty on the number of events that should be eliminated. Consequently, the threshold ratio of the recovered signal to counting error ratio, $S/\sqrt{N}$, may be required to give confidence in the recovery even in the presence of errors in the noise forward model. For example, in the analysis of Gershman et al.[3], Zurbuchen et al.[2], and Raines et al.[5], the threshold used for solar wind heavy ions was $S/\sqrt{N} > 1$ for quiet times and $S/\sqrt{N} > 2$ for active times. The threshold in the study by Gershman et al.[4] for He^+ pickup ions was $S/\sqrt{N} > 1$ and $S/\sqrt{N} > 10$ for quiet and active periods, respectively, to

23

ensure that times of increased solar wind proton flux into the instrument did not lead to an erroneously high measurement of He^+ events.

1. Recovery of individual ion events

The output of the noise processing algorithm is a one-dimensional time history of ion events as a function of E/q. However, to enable the most flexibility in scientific analysis of measured instrument data, a list of real individual events in the full five dimensions is desired. Surprisingly, the low measurement statistics are actually advantageous for this recovery. The processed dataset will have at most a few measured events (often only one) per time step. For this case, there is very little ambiguity on which event per step should be identified as noise. However, in cases with multiple events per time step, a simple criterion is used in which the proper number of events with the highest predicted MCP $x - y$ noise for that particular step are removed. Such a recovery will result in increased error, but this ambiguity will not be common for low statistics, and in fact datasets with large numbers of events will most likely not require such a recovery. Furthermore, the effects of any errors introduced by incorrect assignment of events are expected to be small, especially since most noise events have MCP distributions similar to those of actual events.

As an example, Figure 11 shows the total modeled MCP $x - y$ noise distribution for He^{2+}, He^+, Na^+, and Ca^+, and the corresponding MCP $x - y$ distributions of the recovered noise events, $f^*_{m,noise}$. The overall recovery errors are expected to be less than 25% as predicted by the example distributions from Appendix B. The removed noise events exhibit a similar distribution to the predicted noise in terms of both spatial distributions and intensity. Therefore, despite the fact that the MCP $x - y$ distribution was not considered during the removal process, a reasonable five-dimensional event distribution recovery has been accomplished.

V. CONCLUSIONS

We have developed a mathematical framework for the characterization, forward modeling, and post-processing of noise events in spaceborne time-of-flight mass spectrometers. Using in-flight data from MESSENGER/FIPS as examples, a forward model of sensor noise
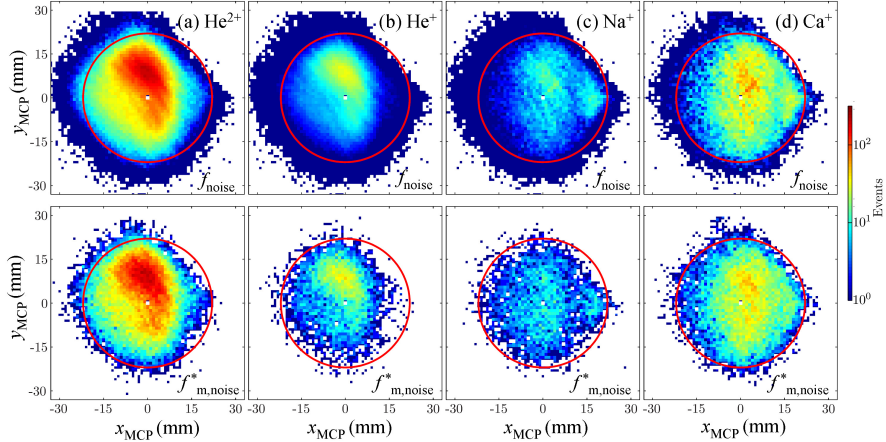
FIG. 11. Total recovered and modeled noise MCP $x - y$ distributions for DOY 100–209, 2011 for (a) He^{2+}, (b) He^+, (c) Na^+, and (d) Ca^+ ions. The recovered noise MCP $x - y$ distributions are the set of individual events (with full five-dimensional information) that have been processed with the Monte Carlo noise removal algorithm. Although the MCP $x - y$ dependence was originally ignored in the event processing, the recovered distribution, $f^*_{m,noise}$, still matches the predicted noise, f_{noise}, implying a quality removal.

was constructed that provided estimations of the ratio of signal to noise in all data dimensions. Kernel density estimation techniques were used to take measured instrument data and recover their underlying distribution functions suitable for use with a Monte Carlo event-processing algorithm. With an accurate noise model, measured events were separated into real and noise-based events, with an expected overall misidentification of events expected to be less than 25%, even for datasets with low SNR or poor counting statistics. This removal was done after accumulating over incident particle directions that were later recovered, enabling maximum versatility in possible scientific analysis.

Although applied to MESSENGER/FIPS as an example, the techniques described here are applicable to any such TOF-MS or even other sensors that measure discrete events. An in-depth characterization of noise not only results in an improved understanding of idiosyncrasies of a particular spaceflight instrument, but can also make any changes in sensor response or newly measured scientific phenomena more easily identifiable. These techniques provide a way to improve the SNR of instruments for which settings may not be sufficiently adjustable in flight, a method to mitigate unexpected instrument behavior in the harsh and unpredictable space environment. Furthermore, existing archived datasets

plagued with low SNR can be revisited and potentially reanalyzed, producing new scientific results from all-but-forgotten instruments.

Appendix A: Event processing errors

Before being applied to in-flight data, the performance of the removal process must be evaluated with synthetic datasets. A direct calculation of the change in SNR after processing data is not necessarily an adequate metric in determining the accuracy of the noise removal. In fact, the results of such a computation may be misleading as, in addition to removing noise-based events, an algorithm may remove some real events, biasing measurements of SNR. Therefore, the relative loss of signal, $|\Delta S/S|$, and relative reduction in noise, $|\Delta\eta/\eta|$, should be analyzed individually. Note that with the presented technique, the total number of recovered real events will be set by the total amount of modeled noise, so errors in processing will likely be a misclassification of noise and real events rather than removal of the incorrect number of events. Therefore, it is expected that $|S| \approx \eta - |\Delta\eta|$.

Steps that contribute to data loss are those for which the total remaining signal ($f^*_{\mathrm{m,real}}$) is less than that of the distribution of true real events ($f_{\mathrm{m,real}}$). Accumulating these differences and dividing by the total number of real events yields the fraction of events lost due to the removal process, $|\Delta S/S|$, as shown in equation (A1),

$$\left|\frac{\Delta S}{S}\right| = \sum_i \begin{cases} \dfrac{f_{\mathrm{m,real}}(x_\mathrm{i}) - f^*_{\mathrm{m,real}}(x_\mathrm{i})}{\displaystyle\sum_j f_{\mathrm{m,real}}(x_\mathrm{j})}, & f_{\mathrm{m,real}}(x_\mathrm{i}) \geq f^*_{\mathrm{m,real}}(x_\mathrm{i}) \\[2em] 0, & f_{\mathrm{m,real}}(x_\mathrm{i}) < f^*_{\mathrm{m,real}}(x_\mathrm{i}) \end{cases} . \tag{A1}$$

On the contrary, steps that overestimate the real signal, i.e., $f^*_{\mathrm{m,real}} > f_{\mathrm{m,real}}$, can be used according to equation (A2) to calculate the fraction of the noise removed, $|\Delta\eta/\eta|$,

$$\left|\frac{\Delta\eta}{\eta}\right| = 1 - \sum_i \begin{cases} \dfrac{f^*_{\mathrm{m,real}}(x_\mathrm{i}) - f_{\mathrm{m,real}}(x_\mathrm{i})}{\displaystyle\sum_j f_{\mathrm{m,noise}}(x_\mathrm{j})}, & f^*_{\mathrm{m,real}}(x_\mathrm{i}) \geq f_{\mathrm{m,real}}(x_\mathrm{i}) \\[2em] 0, & f^*_{\mathrm{m,real}}(x_\mathrm{i}) < f_{\mathrm{m,real}}(x_\mathrm{i}) \end{cases} . \tag{A2}$$

If $f_{\mathrm{m,real}}$ and $f_{\mathrm{m,noise}}$ are known, equations (A1) and (A2) can be used to evaluate the accuracy of a noise removal process. Note that these quantities are not directly calculable

from data and can be used in the analysis only of synthetic datasets. Ideally, $|\Delta S/S| \to 0$ and $|\Delta \eta/\eta| \to 1$, i.e., all of the noise is removed without eliminating any of the true signal.

The overall misidentification of events ($|\Delta f/f|$) can be calculated by adding the number of removed signal events (ΔS) to the number of remaining noise events, ($\eta - \Delta \eta$), and normalizing by the total number of events, N,

$$\left| \frac{\Delta f}{f} \right| = \frac{S + (\eta - \Delta \eta)}{N}. \tag{A3}$$

From the known relationships between S, η, and SNR_{m},

$$S = N \cdot \frac{SNR_{\mathrm{m}}}{1 + SNR_{\mathrm{m}}} \text{ and}$$

$$\eta = N \cdot \frac{1}{1 + SNR_{\mathrm{m}}}, \tag{A4}$$

$$\tag{A5}$$

and combining these relations with equations (A1) and (A2), equation (A3) can be written as,

$$\left| \frac{\Delta f}{f} \right| = \frac{1}{1 + SNR_{\mathrm{m}}} \left(1 - \left| \frac{\Delta \eta}{\eta} \right| \right) + \frac{SNR_{\mathrm{m}}}{1 + SNR_{\mathrm{m}}} \left(\left| \frac{\Delta S}{S} \right| \right). \tag{A6}$$

As expected, for the ideal recovery in which $|\Delta S/S| \to 0$ and $|\Delta \eta/\eta| \to 1$, the misclassification of events $|\Delta f/f| \to 0$.

As discussed in Section II A 2, P_{noise} is used instead of $P_{\mathrm{m,noise}}$ as part of an iterative Monte Carlo event processing technique. The error introduced by this approximation can be quantified through Monte Carlo processing of two synthetic datasets. Two test cases are examined: (1) real events uniformly distributed in time, and (2) real events all occurring at a single time step. Noise sampled from the distribution shown in Figure 2 is superimposed on each signal.

In the case of a perfectly localized signal, $P_{\mathrm{noise}} = P_{\mathrm{m,noise}}$. Therefore, even for small N, a Monte Carlo processing technique will separate $f_{\mathrm{m,real}}$ and $f_{\mathrm{m,noise}}$ with almost 100% accuracy, with $|\Delta S/S| \approx 0$ and $|\Delta \eta/\eta| \approx 1$. For a uniformly distributed signal, however, $P_{\mathrm{noise}} \neq P_{\mathrm{m,noise}}$, especially for small N, resulting in errors in the noise removal process that are dependent on both N and SNR_{m}. For datasets with various SNR_{m}, $|\Delta S/S|$, $|\Delta \eta/\eta|$,
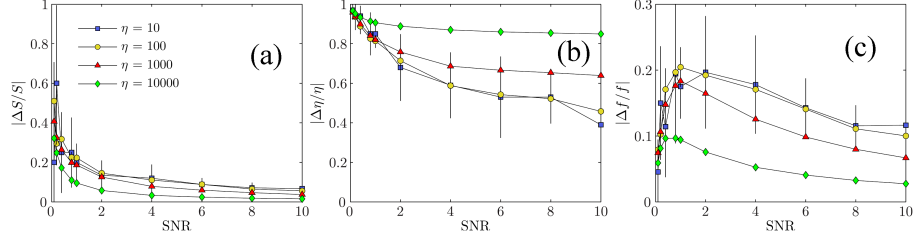
FIG. 12. (a) $|\Delta S/S|$, (b) $|\Delta\eta/\eta|$, and (c) $|\Delta f/f|$ as functions of SNR of a uniformly distributed real event signal for various numbers of noise-based events η. An idealized Monte Carlo event processing algorithm was used with P_{noise} computed using the noise distribution from Figure 2. Each data point represents the average of 10 independent removals of the same dataset, and the error bars indicate the standard deviation. The overall misidentification of events, $|\Delta f/f|$, is less than 20% for all SNR.

and $|\Delta f/f|$ may be computed for a varying number of noise-based events $\eta = 10, 100, 1000$, and 10,000 and are shown in Figure 12.

Both fewer real and noise events are removed with increasing SNR, with approximately a 20% removal of real events and 80% removal of noise events at SNR ≈ 1 in Figures 12a and 12b. The overall event identification error, $|\Delta f/f|$, in Figure 12c, however, is less than 20% for all SNR, peaking near SNR ≈ 1. Event misidentification is a consequence of using the large N limit noise probability P_{noise} instead of $P_{\text{m,noise}}$. These errors will be greatest for non-localized event distributions. As expected, improved results are obtained for large N, where $P_{\text{m,noise}} \to P_{\text{noise}}$. However, despite these misidentifications, it is important to remember that the algorithm will necessarily remove the correct amount of noise, i.e., the total amount of recovered signal will be accurate to within the errors of f_{noise}.

Appendix B: Event processing errors from KDE

With appropriately defined metrics $|\Delta S/S|, |\Delta\eta/\eta|$, and $|\Delta f/f|$, the performance of a KDE-implemented noise removal algorithm can be analyzed using synthetic data as a function of the total number of events and dataset SNR. The two test cases used in Appendix A are re-examined here. Instead of a known P_{noise}, however, an approximate P_{noise} is obtained through KDE estimation of $f(x)$. $|\Delta S/S|$, $|\Delta\eta/\eta|$, and $|\Delta f/f|$ are shown as functions of SNR for both the uniform real event and localized real event distributions in Figure 13.
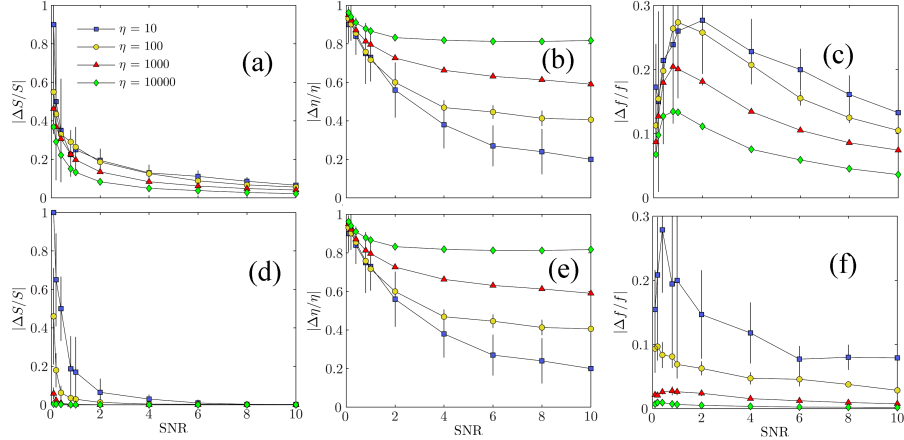
FIG. 13. (a), (b), (c) are the same as Figure 3 but with $f_{\rm m,KDE}$ instead of a known f for a uniform distribution of real events. (d),(e),(f) are the same as parts (a),(b), and (c) but for a localized event source at time step $t = 2800$. With KDE techniques, the overall misidentification error is expected to be less than 25%.

For the uniform real event distribution, the $|\Delta S/S|,|\Delta\eta/\eta|$, and $|\Delta f/f|$ values here are similar to the ideal Monte Carlo process results from Figure 12, with slightly more real events being misidentified at SNR $= 1$ ($\sim$25%) and slightly less noise events removed at higher SNR values ($\sim$20%). The localized events also show some increased error when compared with the idealized case. As expected, an increased number of events N leads to more accurate removal. The KDE technique therefore appears to provide suitable estimates of $f(x)$, such that the end result of the Monte Carlo processing algorithm is similar to that of the ideal cases in all SNR regimes. From these tests, one can expect that for an accurate $f_{\rm noise}$, the total number of expected noise events will be removed with less than 25% of the total events (noise and real) being misidentified for both localized and more uniform distributions of real events.

ACKNOWLEDGMENTS

like to thank Sean C. Solomon for his assistance with the preparation of this manuscript.

REFERENCES

[1]J. A. Gilbert, D. J. Gershman, G. Gloeckler, R. A. Lundgren, T. H. Zurbuchen, T. M. Orlando, J. McLain, R. von Steiger, Characterization of background noise in space-based time-of-flight sensors, Rev. Sci. Instrum., (2013), in revision.

[2]T. H. Zurbuchen et al., Science, **333**, 1862–1865, (2011), doi:10.1126/science.1211302.

[3]D. J. Gershman, T. H. Zurbuchen, L. A. Fisk, J. A. Gilbert, J. M. Raines, B. J. Anderson, C. W. Smith, H. Korth, S. C. Solomon, J. Geophys. Res., **117**, A00M02, 14 pp, (2012), doi:10.1029/2012JA017829.

[4]D. J. Gershman, G. Gloeckler, J. A. Gilbert, J. M. Raines, L. A. Fisk, S. C. Solomon, E. C. Stone, T. H. Zurbuchen, J. Geophys. Res. Space Physics, **118**, 1389–1402, (2013), doi:10.1002/jgra.5022.

[5]J. M. Raines et al., J. Geophys. Res. Space Physics, **118**, 1604–1619, (2012), doi:10.1029/2012JA018073.

[6]T. H. Zurbuchen, G. Gloeckler, J. C. Cain, S. E. Lasley, W. Shanks, in *Conference on Missions to the Sun II*, ed. by C. M. Korendyke, Proceedings of the Society of Photo-Optical Instrumentation Engineers, Bellingham, Wash., vol. 3442, 1998, pp. 217–224.

[7]G. B. Andrews, et. al., Space Sci. Rev., **131**, 523–556 (2007) doi:10.1007/s11214-007-9272-5.

[8]S. C. Solomon et al., Planet Space Sci., **49**, 1445–1465, (2001), doi:10.1016/S0032-0633(01)00085-X.

[9]V. P. Tuzlukov, Signal Processing Noise, 1st ed. (CRC Press, Boca Raton, FL, 2002), 688 pp.

[10]H. Kuo, White Noise Distribution Theory, 1st ed. (CRC Press, Boca Raton, FL, 1996), 400 pp.

[11]S. V. Vaseghi, Advanced Digital Signal Processing and Noise Reduction, 4th ed. (John Wiley, Hoboken, NJ, 2009), 544 pp.

[12]R. C. González and R. E. Woods, Digital Image Processing, 3rd ed. (Prentice Hall, Upper Saddle River, NJ, 2007), 976 pp.

[13]P. Comon and C. Jutten, Handbook of Blind Source Separation: Independent Component

Analysis and Applications, 1st ed. (Academic Press, Waltham, MA, 2010), 856 pp.

[14] A. Smirnov, Processing of Multidimensional Signals, 1st ed. (Springer-Verlag, Berlin, 1999), 284 pp.

[15] G. M. Davis, Noise Reduction in Speech Applications, 1st ed. (CRC Press, Boca Raton, FL, 2002), 432 pp.

[16] M. H. Kalos and P. A. Whitlock, Monte Carlo Methods, 2nd ed. (John Wiley, Hoboken, NJ, 2008), 215 pp.

[17] K. Binder and D. W. Heerman, Monte Carlo Simulation in Statistical Physics: An Introduction, 5th ed. (Springer, New York, 2010), 214 pp.

[18] B. F. J. Manly, Randomization, Bootstrap and Monte Carlo Methods in Biology, 3rd ed. (CRC Press, Boca Raton, FL, 2006), 480 pp.

[19] P. Glasserman, Monte Carlo Methods in Financial Engineering, 1st ed. (Springer, New York, 2003), 616 pp.

[20] D. T. Young, et al., Space Sci. Rev., **114**, 1–112 (2004), doi:10.1007/s11214-004-1406-4.

[21] D. W. Scott, Multivariate Density Estimation: Theory, Practice, and Visualization, 1st ed. (John Wiley, Hoboken, NJ, 1992), 376 pp.

[22] B. W. Silverman, Density Estimation for Statistics and Data Analysis, 1st ed. (CRC Press, Boca Raton, FL, 1986), 176 pp.

[23] R. L. Graham, D. E. Knuth, O. Patashnik, Concrete Mathematics: a Foundation for Computer Science, 2nd ed. (Addison-Wesley, Boston, MA, 1994), 657 pp.

[24] F. Allegrini, D. J. McComas, D. T. Young, J. Berthelier, J. Covinhes, J. Illiano, J. Riou, H. O. Funsten, R. W. Harper, Rev. Sci. Instrum., **77**, 044501, 7pp, (2006) doi:10.1063/1.2185490.

[25] T. H. Zurbuchen et al., Science, **325**, 90–92, (2008), doi:10.1126/science.1159314.