Adv. Space Res. Vol. 13, No. 6, pp. (6)131–(6)139, 1993
Printed in Great Britain.

0273–1177/93 $24.00

VIEWING THE OUTER HELIOSPHERE IN ENERGETIC NEUTRAL ATOMS

K. C. Hsieh* and M. A. Gruntman**

* *Department of Physics, University of Arizona, Tucson, AZ 85721, U.S.A.*
** *Space Sciences Center, University of Southern California, Los Angeles, CA 90089, U.S.A.*

The authors wish to dedicate this article
to the memory of Vladas B. Leonas

ABSTRACT

The outer heliosphere is inaccessible for repeated *in situ* investigations. One can, however, sample the charged-particle populations of that remote region from the inner solar system, including from Earth's orbit, by detecting the energetic neutral atoms produced by charge exchange between the energetic ions and the ambient neutral atoms. Fluxes of energetic neutral atoms emanating from the heliospheric interface region carry information on the thermalization of the solar-wind at the termination shock and the unmodulated anomalous cosmic rays, believed to be accelerated by the termination shock. We examine the techniques for viewing the outer heliosphere in energetic neutral atoms from within.

INTRODUCTION

As the four spacecraft Pioneers 10 and 11 and Voyagers 1 and 2 continue their journey away from our Sun, the anticipation of their crossings of the solar-wind termination shock and the heliopause heightens. Concurrently, as the telemetry coverage shrinks due to budget constraints, the chance of missing the crossings also increases /1/. With the best of luck, there will be at most four point-measurements of the boundary of the heliosphere. The inaccessibility of the outer heliosphere denies our immediate and repeated investigation of that region which withholds answers to a host of questions; for example: What happens to the interplanetary shocks and corotating interaction regions as they propagate into the outer heliosphere? What is the morphology of the merging interaction regions? What happens to the solar wind at and beyond the termination shock, if the shock exists at all? What happens to the cosmic ray particles there? What is the unmodulated galactic cosmic-ray spectrum? Are anomalous cosmic rays produced there? How efficient is the shock acceleration mechanism? How does the particle population there change with solar cycle? *etc.*

The lack of opportunity to answer these questions by means of *in situ* measurements need not result in despair, if we can find means to continuously sense the outer heliosphere remotely from the inner solar system. The 2-3 kHz radiation that could be coming from the solar-wind termination shock is obscured by solar-wind trapping and the possible emissions from Jupiter /2/; the major plasma constituent, protons, cannot be imaged optically; and all charged particles coming from the outer heliosphere would be modulated by the solar wind. The choice of messengers carrying information from the outer heliosphere to us in the inner heliosphere with least possible disturbance narrows down to the energetic neutral atoms (ENA's) /3-5/.

We describe how ENA's can play the role of the emissaries from the outer heliosphere and how they can be detected even under stringent observational conditions. Several reviews are available on ENA techniques with an emphasis on the study of planetary magnetospheres, *e.g.* /6,7/. We accentuate here on the case of remote sensing the outer heliosphere.

ENA, THE MESSENGER

When an energetic ion population shares the same space with a neutral gas, some of the singly ionized ions are neutralized by charge exchange with the neutral atoms of the gas. A singly charged ion spirals around a local magnetic field line until it takes up an electron from an ambient neutral atom to become an ENA. The charge-exchange collision involves little momentum and energy transfer between the two participants. The resulting ENA is freed from the magnetic field and flies away with the instantaneous velocity of the ion at the moment of charge-exchange. The ability to fly across magnetic field lines makes ENA's valued messengers, being a sample of the parent ion population in energy and composition as well as carrying directional information to distant observers. The overwhelming majority of ENA's are hydrogen atoms, henceforth we disregard all minor components.

ENA fluxes come from different ion populations, having different flux levels, and energy-, spatial- and temporal-dependencies. The expansion and the eventual termination of the solar-wind flow separate the ENA-producing ion populations of interest into two groups with respect to the demarcating termination shock: those inside the heliosphere, *i.e.* in the preshocked solar wind, and those at and beyond the shock. In the former, ignoring local emissions from planetary magnetospheres, there are three distinct populations of energetic interplanetary ions whose energies are well above those of the solar wind: the quiet-time interplanetary ions (QTIP), the ions associated with corotating interaction regions (CIR), and the energetic solar particles (ESP). In the latter, there are protons of the shocked or thermalized solar wind (TSW) with energy in the range 0.1 - 2.0 keV, the galactic cosmic rays, and the anomalous cosmic rays (ACR's). Galactic cosmic rays can be neglected here because at the energies where charge exchange is significant (< 1 MeV) the ACR's dominate /8/. These ion populations and the heliosphere are schematically represented in Figure 1. The neutral gas necessary for the production of ENA's is provided by the local interstellar medium (LISM), which penetrates the heliosphere /9,10/. The spatial dependence of both neutrals and ions in and out of the heliosphere dictates the morphologies of the respective ENA fluxes.

ENA's produced at point Q in the outer heliosphere reach an observer A in the inner heliosphere after traversing a distance s (Figure 2). The ENA flux, j_{Ai}, in (cm^{-2} sr^{-1} s^{-1} keV^{-1}), of a given species *i* reaching A is

$$j_{Ai}(\mathbf{R}_0,s,E) = \int_s j_i(\mathbf{R},E) \sum_k [\sigma_{ik}(E)\, n_k(\mathbf{R})]\, \exp[-D(s,E)]\, ds \qquad (1)$$

Here, $j_i(\mathbf{R},E)$ is the directional differential flux of the parent ions, $n_k(R)$ is the number density of species *k* of the ambient gas, $\sigma_{ik}(E)$ is the cross sections for charge exchange between ions of species *i* and neutral atoms of species *k*. The extinction of ENA's between points Q and A is described by

$$D(s,E) = \int_s (\beta_p + \beta_e + \beta_\gamma)\, dt \qquad (2)$$

where the processes include charge exchange with solar wind protons and impact ionization by protons (rate β_p), electron-impact ionization (β_e), and photo-ionization by solar radiation (β_γ) and integration is performed along ENA trajectory s. Collisional ionization rates are determined by proton and electron number densities and relevant cross sections. The latter are shown for hydrogen in Figure 3, where the solid lines are for ENA production while the dashed lines are for ENA removal. The energy dependence of the production cross sections and the usual ion spectra found in interplanetary space suggest that the ENA flux would drop sharply with increasing energy.

The ENA flux, in (cm^{-2} sr^{-1} s^{-1} keV^{-1}), reaching A is independent of the distance between the source and observer A, but it is sensitive to column density, proportional to $j_i(\mathbf{R},E)$ $n_k(\mathbf{R})$ and the thickness

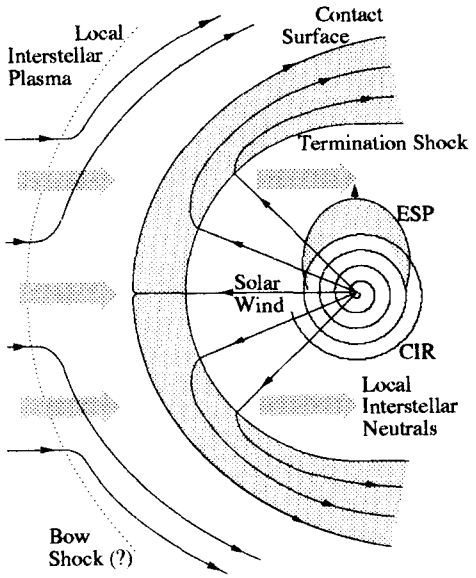

Fig.1. Schematic diagram of the heliosphere.
ESP: energetic solar particles;
CIR: corotating interaction regions.

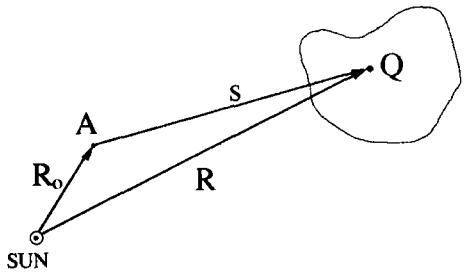

Fig.2. The geometry of remote sensing an ion populations in ENA.

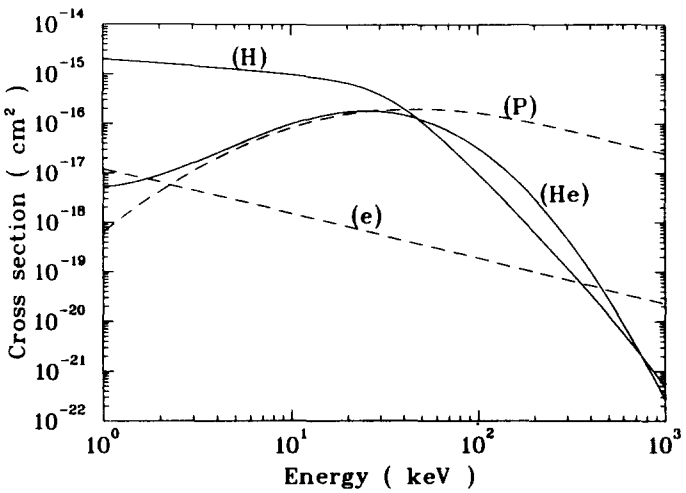

Fig.3. Energy dependence of ENA (hydrogen) production (solid lines) and destruction (dashed lines) cross-sections.
(H): charge exchange with atomic hydrogen;
(He): charge exchange with atomic helium;
(P): proton impact ionization;
(e): electron impact ionization.

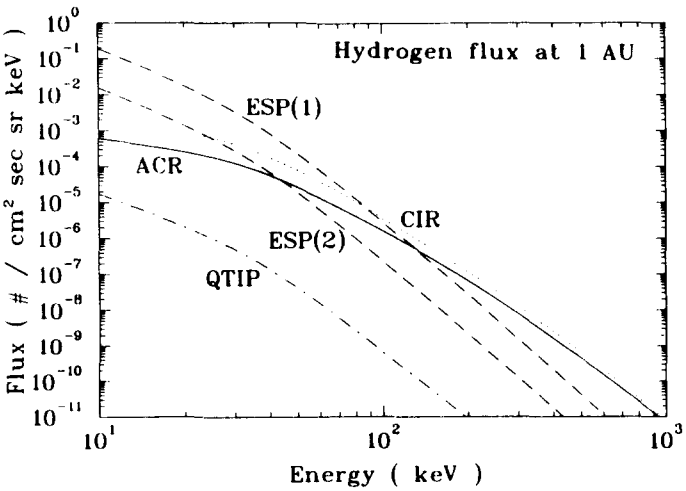

Fig.4. Representative maximum flux (*i.e.* from the direction where it is brightest) spectra of ENA's of different origins seen at 1 AU from the Sun. ESP(1) and ESP(2) correspond to large and small plumes /3/. ENA's of CIR origin have energy spectrum comparable to those of ACR origin, but these two types have distinct spatial and temporal dependencies.

of the production region along the line of sight. Based on the knowledge of $n_k(\mathbf{R})$ /9,10/, one could then deduce $j_i(\mathbf{R},E)$ in the inaccessible outer heliosphere. The techniques of extracting $j_i(R, E)$ from the detected ENA flux $j_A(\mathbf{R_0},s,E)$ are similar to those being developed for planetary magnetosphere imaging (*e.g.* /11,12/).

ENA's produced by the TSW protons beyond the termination shock are relatively low energy, 100-1000 eV, and their flux is highly anisotropic /5/. Simulated spectra of the ENA's of higher energies originating from various ion populations mentioned earlier are shown in Figure 4 /3/. These fluxes display distinctively different characteristics; for example, the ENA flux of ACR origin produced at and beyond the termination shock is less sensitive to solar activities, is characterized by more stable temporal and spatial distributions, and it exceeds the ENA fluxes from CIR at heliocentric distance > 7 AU, and those of ESP origin at > 17 AU. These features, though obviously model dependent, provide us a necessary guide in determining the requirements to techniques for the detection of the heliospheric ENA fluxes.

DETECTION OF ENA's COMING FROM THE OUTER HELIOSPHERE

The relatively low ENA flux reaching an ENA sensor must be sorted out from the usually more intense *in situ* ion fluxes of the same energy and mass and background EUV/UV radiation, especially in hydrogen Ly-α 1216 Å line with intensity 500—1000 R [1 R = 1 Rayleigh = $10^6/4\pi$ photons cm^{-2} sr^{-1} s^{-1})]. Consequently, effective rejection of ions and EUV photons is a crucial requirement for any ENA sensor.

The use of electrostatic or magnetic field to deflect ions from their entrance into the instrument is obvious. Electrostatic field is preferred because 1) magnet designs are typically more complicated than that for an electrostatic device and 2) an electrostatic device has the option of being turned off when desired, thus allowing the same sensor to be used for *in situ* ion investigation. The simplest arrangement of the electrostatic deflector is a pair of parallel plates of length, L, separation, d, and having a potential difference, V. Ignoring fringe fields, a simple trajectory under constant force allows one to approximate the maximum energy E_{max} of ions of charge q such a collimator can reject:

$$E_{max} = (qV/4)\,(L/d)^2 \tag{3}$$

A whole stack of these pairs can certainly multiply the effective geometrical factor, thus capturing more of the rare ENA's. For a given voltage source, maximizing L/d ratio obviously maximizes E_{max}. Such an optimization, however, tends to minimize the instrument field of view (FOV) defined by d/L. Clearly, setting a limit on E_{max} according to the expected ENA spectrum predicted by a realistic model is necessary. Judging from the relevant cross sections (Fig.3), the energy E_{max} = 100 keV for protons is reasonable. A FOV can be increased by creating a tilt between the plate-pair, *i.e.* further apart at the entrance and closer at the exit. Each pair would then have a larger FOV than in the case of parallel plates and have a look direction slightly different from its adjacent pairs, thus turning the stack of plates into a fan-shaped collimator with a spread out FOV. Trajectories of incident ions can be simulated on the computer and verified by calibration (*e.g.*,/13/).

Collimator itself can be a source of noise and other problems. The first is the forward scattering of incident particles by the collimator plates. The solution for similar situations has been the serration of the plate surface. The second, a more subtle one, is the production of photon-electron cascades. When a secondary electron, whether produced by the impact of an ion or a photon, is released from the negatively biased plate, it is accelerated towards the opposite plate. If the electrical potential difference between the plates, V, is sufficiently high, the impact of the accelerated electron could produce an X-ray or EUV/UV photon, which, in turn, has a good chance of striking the negatively biased plate and producing one or more photoelectrons, resulting in the development of an avalanche. Such a electron-multiplying cascade can cause two problems: noise in the detectors due to scattering of these energetic electrons and photons and overloading the power supply of the collimator. The solution calls for the treatment of the plate surfaces to reduce the yield of photoelectrons and the lowering of the voltage , V, to reduce the yield of X-rays. Since the most intense photon background

in space is Ly-α line and since the yield of X-rays requires massive nuclei, treating the plate surfaces with graphite and a limit on the potential difference to 20 kV have been effective /14/. The replacement of graphite-treated aluminum plates with carbon-based composite material is being tested for the Magnetosphere Imaging Instrument (MIMI) to be flown on Cassini. This new approach not only would solve the problem just mentioned, but also should drastically reduce the weight of the ion-rejecting collimator.

Unlike ions, photons cannot be manipulated without directly affecting the incident ENA's, because both photons and ENA's have rectilinear trajectories, thus sharing the same path way to the detector. There are several ways to solve this problem, which, however, would infringe upon the way of analyzing the mass and energy of each incident ENA. We will combine the discussions on EUV/UV rejection and particle identification. The use of submicron diffraction filters to directly block background photons and allow unobstructed passage of ENA's is still a fledgling approach /15/.

If all unwanted ions are rejected by a collimator, then many ion detectors and analyzers can be used for detecting and analyzing ENA's. In the energy range 10 to 200 keV, time-of-flight (TOF) analyzers on the basis of microchannel plate (MCP) detectors and solid-state detectors (SSD's) are widely used /13,14,16,17/. Each incident particle is identified by its TOF over a known distance. A START pulse is generated by a MCP detector when struck by secondary electrons emitted from a foil caused by the passage of the particle. Such a technique to generate a START signal is widely used in nuclear physics experiments and highly transparent and isochronous electron mirrors have been flown on space missions (*e.g.*, /17/). The STOP pulse is generated by another MCP detector that intercepts the incident particle some distance down stream from the first foil. A SSD can be used as a STOP detector and in this case the STOP signal is produced by another detector that receives secondary electrons ejected from the SSD sensitive surface. In such analyzers, there can be triple coincidence signals associated with each particle. The SSD signal provides a measure of the total residual energy of the particle. Combining velocity and energy measurements, the particle's mass is established. Energy threshold of the technique is determined by the energy threshold of SSD, which is typically 10-15 keV/nucleon, and energy losses in the thin foil at the instrument's entrance.

TOF analyzers' response to EUV photons is quite different. The passage of a photon may or may not yield a photoelectron from the foil. If it does yield an electron, then the photon is totally absorbed by the foil and the STOP detector is not triggered. In either case, each photon can generate at most one signal, a START or a STOP. Hence, no photon can by itself constitute an acceptable coincidence event in a TOF analyzer. Photoelectron yields from the foils are well known /18/. When the photon intensity is high, however, it is possible for two photons to randomly mask as one ENA: an earlier one triggers the START detector and a later one triggers the STOP detector. The frequency of such accidental coincidences, f, can be estimated for any given TOF analyzer defined by its coincidence gate, t, *i.e.* the maximum TOF expected for the intended incident particles, and the individual counts in each detector, C_i. For the case of only one STOP signal, $f = t\, C_1\, C_2$. For example, for $C_1 = C_2 = 10^4\ s^{-1}$ and t= 100 ns, the random coincidence rate would be $f = 10\ s^{-1}$. Individual detector count rates, C_i, must be reduced by the thin foil at the instrument entrance /18,19/. Penetrating particles, such as cosmic rays, pose no threat. In the case of TOF analyzers, penetrating cosmic-ray fluxes being orders of magnitude lower than Ly-α fluxes, the coincidence requirement of TOF more than adequately provides the immunity. In the case of SSD, selective pulse-height windows and anti-coincidence detectors can also effectively reject such contaminations.

An example of the TOF instrument with the SSD detector is shown in Figure 5. This instrument, HSTOF, is under preparation for the SOHO mission and will be capable of measuring ENA's in the L1 point of the Sun-Earth system. In order to reduce energy threshold of the instruments with the SSD detectors, it was suggested /20/ to measure the number of secondary electrons emitted from the thin foil during the particle passage through it. Statistics of secondary electron emission from the thin foils /20/ determines the efficiency of such a technique. This approach is utilized for the first time in the instrument Magnetosphere Imaging Instrument (MIMI) which is being developed for the Cassini mission (Figure 6). This experiment will measure heliospheric ENA's on the spacecraft's way to Saturn.

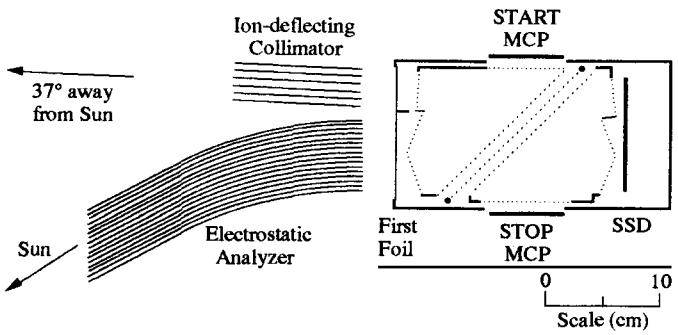

Fig.5. ENA instrument HSTOF of the SOHO mission. Field of view 2°×34°. (Courtesy of D.Hovestadt)

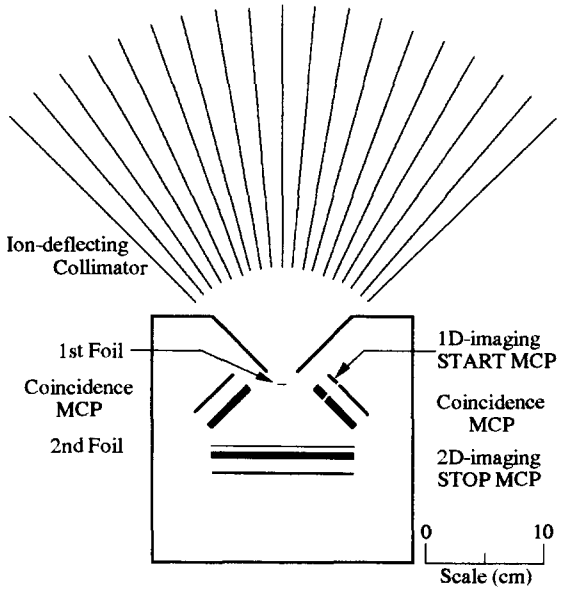

Fig.6. ENA instrument MIMI of the Cassini mission. Field of view is 90°×120°; STOP MCP detector sensitive area is 10×10 cm. Pulse height analysis of the STOP signal provides capability to distinguish between H and O. (Courtesy of D.G.Mitchell)

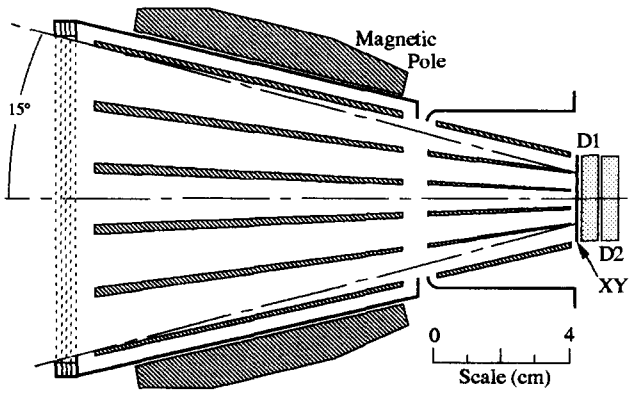

Fig.7. Imaging ENA instrument based on the solid-state detector similar to that described in /21/.

Obviously, TOF analyzers are extremely effective in rejecting EUV-initiated noise and can provide both the velocity and energy analysis when a SSD is used. This approach, however, does have one worrisome feature, *i.e.* the need of very thin foils as both EUV attenuators and START-signal generators. The transmittance of Ly-α in different materials have been measured, *e.g.* /18,19/. The requirement for high attenuation for EUV photons and that for low stopping-power for particles work against each other. Typically, these foils are ~ 1000 Å in thickness. In addition, the durability of these foils directly exposed to the EUV/UV illumination are more of a worry for ENA instruments than in previously flown instruments where thin foils are used only as START-signal generators situated inside a spectrometer.

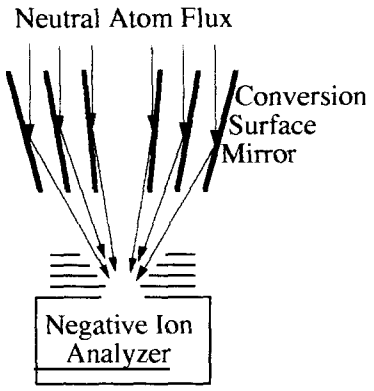

Fig.8. Schematic view of the detector to measure "low"-energy ENA from heliospheric interface based on novel conversion surfaces.

Along a different path, Voss et al. /21/ have been flying SSD-based spectrometers on satellites such as CRRES. These detectors have large sensitive area (2.5 x 2.5 cm), necessary for low-intensity ENA fluxes, and they can claim immunity to the intense Ly-α of the geocorona. The SSD has a 20 μm thick aluminum protective layer and each pixel has its own charge-sensitive amplifier and signal-processing logic on an integrated chip with the electronic threshold equivalent to 8 keV in particle energy. Such a device is capable of detecting heliospheric ENA fluxes without any worry of EUV/UV contamination. Since the intensity of photon flux in the heliosphere, when not staring at the Sun, is less than that of the geocorona, the use of a SSD with a moderate aluminum window, small pixels, and micro-electronics, is an alternative to TOF that needs thin foils (Figure 7). Although the instrument of the type /21/ is not capable of mass identification, it can be efficiently used in the heliospheric studies because Mass of ENA's is typically known from physical processes involved in ENA production.

Detection of ENA's produced by the TSW protons beyond the termination shock requires use of different technique. ENA energies, 100—1000 eV, are too small to be detected by the SSD or penetrate 1000 Å thick protective foil at the instrument entrance. Two ways to solve this problem has been found. The first one is based on the use of ultra-thin (< 100 Å) foil at the TOF instrument entrance directly exposed to the EUV/UV background /22/. This foil is used for START signal generation and extraction of the ENA signal from the intense background is performed by stringent coincidence requirements. Instrument based on this principle was developed for the study of the neutral component of the solar wind and ENA fluxes from the heliospheric interface region /23/. The second approach is based on the stripping of ENA's in ultra-thin (< 100 Å) foils followed by ion detection and identification by electrostatic and TOF analyzers /24/.

A novel approach to detect low-energy ENA's is currently being studied, which utilizes the conversion of ENA's to negative ions at the specially prepared surfaces /25/. Possible design of the instrument based on such a conversion is shown in Figure 8, which includes a grazing-incidence-telescope-like structure providing a large sensitive area of the instrument.

ACKNOWLEDGEMENT

The information concerning HSTOF is provided by D.Hovestadt and that on MIMI by D.G.Mitchell. This work was supported partially by NASA grants NAG5-260, NAGW 1419, and NAGW 1676 to the University of Arizona and NAG 2-146 to the University of Southern California.

REFERENCES

1. J.A.Simpson, The heliosphere: a laboratory for investigating fundamental processes in high energy astrophysics, this issue.

2. W.S.Kurth, Low frequency radio emission from the termination shock, this issue.

3. K.C.Hsieh, K.L.Shih, J.R.Jokipii, and S.Grzedzielski, Probing the heliosphere with energetic neutral atoms, *Astrophys. J.* 393, 756 (1992).

4. E.C.Roelof, Imaging heliospheric shocks using energetic neutral atoms, in: *Solar Wind Seven*, eds.E.Marsch and R.Schwenn, Pergamon Press, 385 (1992).

5. M.A.Gruntman, Anisotropy of the energetic neutral atom flux in the heliosphere, *Pl. Sp. Sci.* 40, 439 (1992).

6. R.W.McEntire and D.G.Mitchell, Instrumentation for the imaging of energetic neutral atoms, in: *Solar System Plasma Physics*, ed. J.H.Waite Jr., J.L.Burch and R.L.Moore, AGU, 69 (1989).

7. K.C.Hsieh, C.C.Curtis, C.Y.Fan, and M.A.Gruntman, Techniques for the remote sensing of space plasma in the heliosphere via energetic neutral atoms: a review, in: *Solar Wind Seven*, eds. E.Marsch and R.Schwenn, Pergamon Press, 357 (1992).

8. E.R.Christian, A.C.Cummings and E.C.Stone, Evidence for anomalous cosmic-ray hydrogen, *Astrophys. J* 334, L77 (1988).

9. T.E.Holzer, Neutral hydrogen in interplanetary space, *Rev. Geophys. Space Phys.* 15, 467-489 (1977).

10. V.B.Baranov, Gasdynamics of the solar wind interaction with the interstellar medium, *Space Sci. Rev.* 52, 89 (1990).

11. E.C.Roelof, Energetic neutral atom image of a storm-time ring current, *Geophys. Res. Lett.* 14, 652 (1987).

12. E.C.Roelof, B.H.Mauk, and R.R.Meier, Instrument requirements for imaging the magnetosphere in extreme-ultraviolet and energetic neutral atoms derived from computer-simulated images, in: *Instrumentation for Magnetospheric Imagery, Proc. SPIE 1744*, 19 (1992).

13. E.P.Keath, G.B.Andrews, A.F.Cheng, S.M.Krimigis, B.H.Mauk, D.G.Mitchell, and D.J.Williams, Instrumentation for energetic neutral atom imaging of magnetosphere, in: *Solar System Plasma Physics*, ed. J.H.Waite Jr., J.L.Burch and R.L.Moore, AGU, 165 (1989).

14. G.Gloeckler, Ion composition measurement technique for space plasma, *Rev. Sci. Instrum.* 61, 3613 (1990).

15. M.A.Gruntman, Submicron structures: promising filters in EUV — a review, in: *EUV, X-Ray, and Gamma-Ray Instrumentation for Astronomy II, Proc. SPIE 1549*, 385 (1991).

16. G.Gloeckler and K.C.Hsieh, Time-of-flight technique for particle identification at energies from 2-400 keV/nucleon, *Nucl. Instrum. Meth.* 165, 537 (1979).

17. B.Wilken and W.Stuedemann, A compact time-of-flight mass spectrometer with electrostatic mirrors, *Nucl. Instrum. Meth.* 222, 587 (1984).

18. K.C.Hsieh, E.Keppler, and G.Schmidtke, Extreme ultraviolet induced forward photoemission

from thin carbon foils, *J. Appl. Phys.* 51, 2242 (1980).

19. K.C.Hsieh, B.R.Sandel, V.A.Drake, and R.S.King, H Lyman α transmittance of thin C and Si/C foils for keV particle detectors, *Nucl. Instrum. Meth. Phys. Res. B* 61, 187 (1991).

20. M.A.Gruntman, A.A.Kozochkina, and V.B.Leonas, Multielectron secondary emission from thin foils bombarded by accelerated beams of atoms, *Sov. Phys. - JETP Lett.* 51, 22 (1990).

21. H.D.Voss, E.Herzberg, A.G.Ghielmetti, S.J.Battel, K.L.Alpert, B.R.Higgins, D.O.Murray, and R.R.Vondrak, The medium energy ion mass and neutral atom spectrometer, *J. Spacecraft Rockets* 29, 566 (1992).

22. M.A.Gruntman and V.A.Morozov, H atom detection and energy analysis by use of thin foils and TOF technique, *J. Phys. E* 15, 1356 (1982).

23. M.A.Gruntman, S.Grzedzielski, and V.B.Leonas, Neutral solar wind experiment, in: *Physics of the Outer Heliosphere*, eds. S.Grzedzielski and D.E.Page, Pergamon Press, 355 (1990).

24. D.J.McComas, B.L.Barraclough, R,C.Elphic, H.O.Funsten III, and M.F.Thomsen, Magnetospheric imaging with low-energy neutral atoms, *Proc. Nat. Acad. Sci. U.S.A.* 88, 9598 (1991).

25. M.A. Gruntman, A new way to measure the composition of the interstellar gas surrounding the heliosphere, this issue.