Inhomogeneity in Cosmic Ray Sources as the Origin of the Electron Spectrum and the PAMELA Anomaly

Nir J. Shaviv,[1] Ehud Nakar,[2] and Tsvi Piran[1]

[1]*Racah Institute of Physics, Hebrew University of Jerusalem, Jerusalem 91904, Israel*
[2]*The Raymond and Beverly Sackler School of Physics & Astronomy, Tel-Aviv University, Tel-Aviv 69978, Israel*
(Received 16 April 2009; revised manuscript received 7 July 2009; published 10 September 2009)

We show that inhomogeneity of cosmic ray (CR) sources, due to the concentration of supernova remnants (SNRs) towards the galactic spiral arms, can naturally explain the anomalous increase in the positron/electron ratio observed by PAMELA. We consistently recover the observed positron fraction between 1 and 100 GeV using SNRs as the sole source of CRs. The contribution of a few known nearby SNRs dominates the CR electron spectrum above ~ 100 GeV, leading to the relatively flat spectrum observed by Fermi and to the sharp cutoff observed by H.E.S.S.

PAMELA [1] discovered that the cosmic ray (CR) positron/electron ratio increases with energy above ~ 10 GeV. This ratio should decrease according to the standard scenario, in which CR positrons are secondaries formed by interactions between the primary CR protons and the interstellar medium (ISM) [2–4]. This apparent discrepancy is now commonly known as the "PAMELA anomaly." It is commonly interpreted as evidence for a new source of primary CR positrons: WIMPs [5,6] or pulsars [7–12]. At higher energies, ATIC [13] shows a peak of CR electrons at 600 GeV. However, this excess is not seen by Fermi [2] or H.E.S.S. [3]. At 1–4 TeV, H.E.S.S. measures [3] a sharp decay in the electron spectrum.

In the standard picture, CRs are thought to originate in supernova remnant (SNR) shocks. CRs diffuse within the disk, and escape once they reach the halo height, $l_H \sim 1$ kpc, above the disk. Most approximate the diffusion coefficient as $D = D_0(E/E_0)^\beta$ and assume that CRs are produced with a power-law spectrum, $N_E \equiv dN/dE \propto E^{-\alpha}$ where the value of α appropriate to each species will be indicated by a subscript. The observed spectrum is a convolution of the source spectrum and propagation losses, giving for the primary electrons $\phi^-(E) \propto E^{-(\alpha_e + \beta)}$. Positrons are secondary CRs formed from CR protons, and suffer additional propagation loses, implying $\phi^+(E) \propto \phi_p(E)E^{-\beta} \propto E^{-(\alpha_p + 2\beta)}$, where $\phi^\pm$ and ϕ_p are the CR positrons, electrons and protons observed fluxes. The predicted flux ratio is $\phi^+/(\phi^- + \phi^+) \approx \phi^+/\phi^- \propto E^{\alpha_e - \alpha_p - \beta}$. Both electrons and protons are expected [14] to have similar spectral slopes, i.e., $\alpha_e \approx \alpha_p$, which is somewhat larger than 2 [15]. Consequently, $|\alpha_p - \alpha_e| < \beta \approx 0.3$–0.6 and the standard model predicts, in contrast to PAMELA observations, a decreasing ϕ^+/ϕ^-.

The diffusing electrons and positrons cool via synchrotron and inverse-Compton scattering, with $dE/dt = -bE^2$. This steepens both the electron and positron spectra at an energy where the cooling time equals the typical

electron and positron age. However, since both suffer the same losses, this does not affect ϕ^+/ϕ^-. Additional effects such as spallation and annihilation can be safely ignored at the energies of interest.

The break from a decreasing to an increasing ϕ^+/ϕ^- was most often interpreted as a break in the positron population, leading to models with additional positron sources. Instead of a harder positron component, the anomaly can be explained by a softer electron component [16]. For example, the anomaly can be explained away if the secondary positron spectrum in the energy range of 10–100 GeV can be approximated by a power-law with an index of $\alpha_p + \beta \sim -3.03$, while the electron spectrum is steeper than -3.04 in the same energy range, as seen by FERMI/LAT. This is also the motivation of our work. We show here how such a situation with a steeper electron spectrum can naturally arise once modifying the source distribution to include inhomogeneities.

The standard model assumes a homogenous source distribution [4,17]. However, since in spiral galaxies star formation is concentrated in spiral arms [18,19] and SNRs are the canonical sources of CRs, one should consider the effect of inhomogeneities in the CR source distribution on intermediate scales (i.e., scales smaller than the Galactic size but large enough such that discrete sources do not have a strong effect) on the CR spectrum. This inhomogeneity of sources influences the electron or positron spectra via radiative cooling which sets a typical distance scale that an electron or positron with a given energy can diffuse away from its source. For a homogenous distribution, cooling affects the spectra of (primary) electrons and (secondary) positrons in the same way, and their ratio is unaffected. On the other hand, primary electrons will be strongly affected by an inhomogeneous source distribution at energies for which the diffusion time is longer than the cooling time. Protons are not affected by cooling and are therefore distributed rather smoothly in the galaxy even if their sources are inhomogeneous. The sec-

ondary positrons (that are produced by the smoothly distributed protons) are only weakly affected by the inhomogeneity of the sources. This effect of a steepened electron spectrum compared to the positron spectrum would induce an observed signature on ϕ^+/ϕ^-, with similar properties to the one observed by PAMELA.

Motivated by this expectation, we construct, first, a simple analytic model for diffusion from an inhomogeneous source. Consider a source at a distance d from Earth. We model the galaxy as a two dimensional slab (see Fig. 1). The x coordinate (the Galactic plane) is infinite, and the y coordinate (the disk height) is finite, l_H. The source is at the origin, and Earth is at $(d, 0)$. A CR diffuses within this slab with a constant diffusion coefficient $D(E)$, and it escapes once $|y| > l_H$. The contribution of CR protons that were generated at time t' to the flux at time t_0 can be approximated as [20]

$$\phi_p(d, t') \propto (Dt)^{-1/2} \exp[-(t/\tau_e) - (\tau_d/2t)], \quad (1)$$

where $t \equiv t_0 - t'$, $\tau_e \approx l_H^2/D$ is the typical escape time and $\tau_d \approx d^2/D$ is the typical diffusion time from the source to Earth. t integration for a steady source yields

$$\phi_p(d) \propto \exp[-\sqrt{2\tau_d/\tau_e}]/D, \quad (2)$$

with a similar energy dependence (via D) as for uniformly distributed sources. The average age of an observed proton is $a = l_H(l_H + \sqrt{2}d)/2D \approx \max\{\tau_e, (\tau_e\tau_d)^{1/2}\}$.

We approximate the cooling effect on the electron's flux as $\phi^-(d, t') \propto \phi_p(d, t')\exp[-t/\tau_c]$, where τ_c is the typical cooling time. Integration over t reads

$$\phi^-(d) \propto \exp[-2\sqrt{\tau_d/\tau_c + \tau_d/\tau_e}]/D\sqrt{1 + \tau_e/\tau_c}. \quad (3)$$

If $\tau_c < \min\{\tau_d, (\tau_e\tau_d)^{1/2}\}$, ϕ^- drops exponentially with decreasing τ_c, while for larger τ_c, ϕ^- is proportional to D^{-1} (relative to the source's spectrum). This is different than the case of uniformly distributed sources.

The positron source function is approximately proportional to $\phi_p(d)$. As positrons and electrons have the same cooling rate, a source at x' contributes to the positron flux

at d approximately $\phi^-(x' - d)$. Therefore,

$$\phi^+(d) \propto \int_{-\infty}^{\infty} \phi_p(x')\phi^-(x' - d)dx'$$

$$\propto \frac{\tau_c}{D}\left(\exp\left[-\sqrt{\frac{2\tau_d}{\tau_e}}\right] - \frac{\exp[-\sqrt{\frac{2\tau_d}{\tau_c} + \frac{2\tau_d}{\tau_e}}]}{\sqrt{1 + \tau_e/\tau_c}}\right). \quad (4)$$

For $\tau_c \gg \tau_e$, the energy dependence of ϕ^+ relative to the source spectrum, $\phi_p^{(s)}$, is $\phi^+/\phi_p^{(s)} \propto D^{-2} \propto E^{-2\beta}$ while for $\tau_c \ll \tau_e$, $\phi^+/\phi_p^{(s)} \propto \tau_c/D \propto E^{-\beta-1}$. This behavior is similar to the one from uniformly distributed sources.

Equations (3) and (4) show that for a source at a distance d from Earth, a turnover in ϕ^+/ϕ^- is observed at E_b which satisfies $\tau_c(E_b) \approx \min\{\tau_x(E_b), [\tau_e(E_b)\tau_x(E_b)]^{1/2}\}$. ϕ^+/ϕ^- for $E < E_b$ decreases, while it increases for $E > E_b$. The typical age of CR protons with energy E_b is $a \sim \max\{\tau_e, (\tau_e\tau_d)^{1/2}\}$. Therefore, a natural prediction of the model is $a(E_b) \gtrsim \tau_c(E_b)$. A comparison of the two observables provides a consistency test for the model. Moreover, over a wide range of the parameter space for which $d \gtrsim l_H$, the model predicts $a(E_b) \approx \tau_c(E_b)$ regardless of the value of the diffusion coefficient D.

Electrons and positrons cool as $dE/dt = -bE^2$ where [21] $b \approx 1.8 \times 10^{-16}$ GeV^{-1} s^{-1} at 1 GeV (and $b \approx 1.4 \times 10^{-16}$ GeV^{-1} s^{-1} at 1 TeV), implying a cooling time $\tau_c = 1/(bE) \approx 17$ Myr at $E \approx 10$ GeV. Observational constraints on the typical proton CR age are measured at a few 100 MeV. Typical ages obtained are 18^{+8}_{-9} Myr [22], 27^{+19}_{-9} Myr [23], or 30^{+21}_{-10} Myr [24]. At 10 GeV, the age should be smaller by a factor of $\sim$1–3, depending on the exact energy dependence of the diffusivity. Thus, according to the observations $a(10 \text{ GeV}) \approx \tau_c(10 \text{ GeV}) \approx 10$ Myr. This apparent coincidence which is explained naturally by our model encourages us to look for a dominant CR source at a distance of a $\sim$kpc from earth. Indeed, the nearest spiral arm to Earth, the Sagittarius-Carina arm, is just the distance needed to explain PAMELA's observations.

To demonstrate quantitatively the potential of this model to recover the observed behavior of ϕ^+/ϕ^-, we simulated numerically the CR diffusion for a realistic spiral-arm concentrated source distribution (see also Ref. [19]). Before presenting these results, we stress that most models explaining PAMELA invoke a new source of high energy CR positrons which has a negligible effect on low energy CR components. However, in our model, the explanation is intimately related to low and intermediate energy CR propagation in the Galaxy. Namely, by revising the source distribution of CRs, we affect numerous properties of $\sim$GeV CRs. Given that the interpretation of observations (in particular, isotopic ratios) used to infer model parameters (such as D_0, β or l_H) depend on the complete model, one should proceed while bearing in mind that these parameters may differ in our model from present canonical values. In this sense, the objective of this Letter is not to carry a comprehensive parameter study, fitting the whole
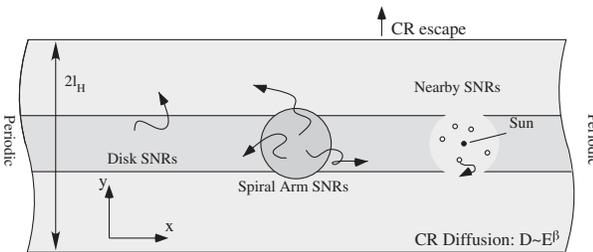

FIG. 1. The galaxy is modeled as a slab of width $2l_H$, with $l_H = 1$ kpc. Beyond $y = \pm l_H$, the CRs escape at a negligible time. CR sources are located in both cylinder shaped arms with a Gaussian cross section of width $\sigma = 300$ pc, and disk sources, with a vertical scale height of 100 pc. The smooth disk distribution is truncated for $r < 0.5$ kpc and $t < 0.5$ Myr.

CR data set to an inhomogeneous source distribution model. Instead, our goal is to demonstrate the potential of the model to explain naturally the PAMELA anomaly.

The geometry of the model is described in Fig. 1. We assume a spiral-arm/disk SNe ratio of 4. The overall normalization of the sources was fit to give the electron spectrum at 10 GeV. The positron production was normalized to give the positron to electron ratio at the same energy. For the ISM density, we took the functional dependence from Ref. [17]. More on the choice of the parameters can be found in Ref. [19].

We take a diffusivity of the form $D = D_0(E/1 \text{ GeV})^\beta$ for $E > 4$ GeV and $D = D_0(4 \text{ GeV}/1 \text{ GeV})^\beta$ for $E < 4$ GeV. Such a break is required to explain the observed break in the CR B/C ratio [17] (though it does not play an important role here). We take $\beta = 1/3$ (corresponding to turbulence with a Kolmogorov spectrum) and $\alpha_e = \alpha_p = 2.37$ such that the predicted proton spectrum will be consistent with the observed proton CR slope of 2.7. We also take $D_0 = 6 \times 10^{27}$ cm^2/sec, which reproduces the break energy in the electron spectrum and the positron fraction. As estimated by the analytic model, the cosmogenic age we obtain in the simulation (14 Myr at 1 GeV per nucleon) is consistent with the observations. Not surprisingly, the halo size and diffusivity considered here are somewhat different (on the low side) relative to standard values often found in the homogeneous model. However, they are within the allowed range (e.g., see [16]). Moreover, as mentioned earlier, the canonical values used in standard diffusion model were obtained by fitting a homogeneous source distribution to the observations, and they therefore get modified under an inhomogeneous source distribution. Note also that while this is the value of D_0 that we found in the specific model that we use, the principle effect of source inhomogeneity on the steepening of the electron spectrum hardly depends on the value of D_0.

Small scale inhomogeneities are important at energies larger than a few hundreds GeV, for which the lifetime, and therefore propagation distance, of electrons is so short that the electron spectrum is dominated by a single, or at most a few nearby sources [9,12,21]. To take this effect into account, we truncate the "homogeneous" disk component at $r < 0.5$ kpc and age less than $t < 0.5$ Myr, and we add all SNRs within this 4 volume: Geminga, Monogem, Vela, Loop I, and the Cygnus Loop, as discrete instantaneous sources. These sources were described using the analytical solution [9] for the diffusion and cooling from an instantaneous point source. For the overall normalization of the point sources, we use the synchrotron observations of SN 1006, which together with the x rays constrain the total energy and magnetic field [25]. In particular, electrons with energy >1 GeV are found to carry $\approx 2 \times 10^{48}$ erg, corresponding to 0.2% out of the total $\sim 10^{51}$ erg mechanical energy in SNRs. For the actual fit in Fig. 2, we assume that all nearby sources are similar and accelerate 1.5×10^{48} erg in electrons. This number, however, is not very

well constrained. Note that due to their very young age, the discrete sources contribute a negligible amount of positrons, nor do they offset the cosmogenic age.

The lower panel of Fig. 2 depicts $\phi^+/(\phi^+ + \phi^-)$ obtained by the simulation. As expected from the simple analytical model, the fraction decreases up to ~ 10 GeV, and then it starts increasing. This explains the so-called PAMELA anomaly. As the CR protons and antiproton spectra are unaffected, our results are consistent with PAMELA's observations of no excess in the antiproton/proton ratio at the same energy range [26]. At about 100 GeV, the ratio flattens, and it decreases above this energy because of the injection of "fresh" CRs from recent nearby SNRs whose high energy primary electrons do not have time to cool.

The upper panel of Fig. 2 depicts the electron spectrum and its constituents—primary spiral-arm electrons, primary disk electrons (without nearby sources), nearby sources, and secondary pairs. There are two minor bumps in the $E^3 N_E$ plot. The lower energy bump arises from spiral-arm electrons, the higher energy of which cannot reach us due to cooling. The higher energy bump is due to a few nearby SNRs. The three "steps" are due to the cooling cutoffs from Geminga, Loop I, and the Monogem SNRs.
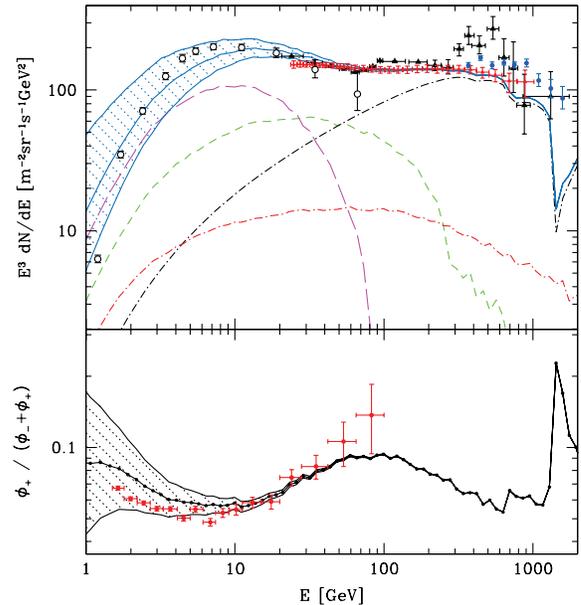

FIG. 2 (color). *Bottom Panel*: Model results and the measured PAMELA points for the positron fraction. The shaded region is the variability expected from solar modulation effects [27]. *Top Panel*: The expected electron and positron spectra—Primary arm electrons (long dashed purple), primary disk electrons with nearby sources excluded (short dashed green), nearby SNRs (dot-dashed black), secondary positrons (dot-dashed red), and their sum (blue). The hatched region describes the solar modulation range (from 200 MV to 1200 MV). The three data sets plotted in black are of HEAT [28] (circles), ATIC [13] (triangles), and H.E.S.S. [3] (open squares). The FERMI [2] results are plotted in red.

The high energy behavior is very sensitive to the exact diffusion model parameters and the poorly constrained SNR energy output in electrons. For the assumed single-SNR normalization choice, a reasonable fit to the FERMI data is obtained. However, a 50% higher normalization would produce a reasonable fit to the ATIC data. In other words, having an ATIC-like peak, or a flatter FERMI-like spectrum, are both possible outcomes of the model.

While the predictions for ϕ^+/ϕ^- for the spiral arms CR model are very different than for a homogenous sources distribution, the effect on the electron spectrum is much more subtle. Both models predict a break of the electron spectrum at 10 GeV. The break predicted by spiral-arm model is from a power law to an exponential, while in the homogenous model it is a broken power law. Given that above ~ 100 GeV the electron spectrum is strongly affected by local SNRs, the energy range between 10 to 100 GeV is too short to distinguish, based on the electron spectrum alone, between the two models. Thus, while both models can adequately reproduce the observed electron spectrum (at least up to 100 GeV), only the inhomogeneous source model can explain the observed ϕ^+/ϕ^-.

One of the interesting predictions of the model where the electrons observed above ~ 100 GeV are dominated by local SNRs is that these sources produce only primary electrons and have only negligible contribution to secondary positron flux. As a result, the positron/electron ratio must decrease at a few hundred GeV, which is not far above the present PAMELA measurement. It should reach a minimum around a TeV, where it should start rising again. Whether or not it can go up to about 50% at a few TeV depends on whether the CRs from very recent SNe, the Cygnus Loop, and Vela could have reached us or not. This critically depends on the exact diffusion coefficient. Here, it is also worth pointing out that above a few TeV, the secondaries must be produced within the local bubble, implying that their normalization should be 10 times lower than for the lower energy secondaries. These predictions are in contrast to the case where positrons above 10 GeV and electrons above 100 GeV are due to a primary source of pairs, in which case the positron fraction is expected to keep rising also at a few hundreds GeV. With these predictions, it will be straightforward in the future to distinguish between propagation induced ''anomalies,'' and real anomalies arising from primary pairs (in particular, when PAMELA's observations will extend to higher energies). Of course, it is possible that a few hundred GeV electrons are due to a source of primary pairs, while the PAMELA anomaly is a result of SNRs in the spiral arms, but then it would force us to abandon the simplicity of the model, that the anomalies are all due to propagation effects from a source distribution borne from the known structure of the Milky Way.

Irrespectively, this work demonstrates that the intermediate scale inhomogeneities expected in the CR source distribution leave nontrivial imprints on the electron and positron spectra. These should be further investigated before reaching definitive conclusions about the existence of primary positron sources.

[1] O. Adriani *et al.*, Nature (London) **458**, 607 (2009).
[2] A. A. Abdo *et al.*, Phys. Rev. Lett. **102**, 181101 (2009).
[3] HESS Collaboration, arXiv:0905.0105.
[4] I. V. Moskalenko and A. W. Strong, Astrophys. J. **493**, 694 (1998).
[5] L. Bergström, T. Bringmann, and J. Edsjö, Phys. Rev. D **78**, 103520 (2008).
[6] A. Ibarra and D. Tran, J. Cosmol. Astropart. Phys. 02 (2009) 021.
[7] A. K. Harding and R. Ramaty, *International Cosmic Ray Conference* (Nauk, Moscow, 1987), Vol. 2, p. 92.
[8] X. Chi, K. S. Cheng, and E. C. M. Young, Astrophys. J. Lett. **459**, L83 (1996).
[9] A. M. Atoyan, F. A. Aharonian, and H. J. Völk, Phys. Rev. D **52**, 3265 (1995).
[10] D. Hooper, P. Blasi, and P. Dario Serpico, J. Cosmol. Astropart. Phys. 01 (2009) 25.
[11] H. Yuksel, M. D. Kistler, and T. Stanev, Phys. Rev. Lett. **103**, 051101 (2009).
[12] S. Profumo, arXiv:0812.4457.
[13] J. Chang *et al.*, Nature (London) **456**, 362 (2008).
[14] R. Blandford and D. Eichler, Phys. Rep. **154**, 1 (1987).
[15] N. Duric *et al.*, Astrophys. J. **445**, 173 (1995).
[16] T. Delahaye *et al.*, Astron. Astrophys. **501**, 821 (2009).
[17] A. W. Strong and I. V. Moskalenko, Astrophys. J. **509**, 212 (1998).
[18] C. K. Lacey and N. Duric, Astrophys. J. **560**, 719 (2001).
[19] N. J. Shaviv, New Astron. Rev. **8**, 39 (2003).
[20] We assume for simplicity that the diffusion is one dimensional. This results with an exponent once integrated. Two dimensional diffusion (from a linear spiral arm) would give a less transparent Bessel function.
[21] T. Kobayashi, Y. Komori, K. Yoshida, and J. Nishimura, Astron. J. **601**, 340 (2004).
[22] M. E. Wiedenbeck and D. E. Greiner, Astrophys. J. Lett. **239**, L139 (1980).
[23] A. Lukasiak, P. Ferrando, F. B. McDonald, and W. R. Webber, Astrophys. J. **423**, 426 (1994).
[24] J. A. Simpson and M. Garcia-Munoz, Space Sci. Rev. **46**, 205 (1988).
[25] T. Yoshida and S. Yanagita, in *The Transparent Universe*, edited by C. Winkler *et al.*, (ESA Special Publication, 1997), Vol. 382, p. 85.
[26] O. Adriani *et al.*, Phys. Rev. Lett. **102**, 051101 (2009).
[27] J. M. Clem *et al.*, Astrophys. J. **464**, 507 (1996).
[28] M. A. DuVernois *et al.*, Astrophys. J. **559**, 296 (2001).