Solar Wind Magnetic Field Correlation Length: Correlation Functions versus Cross-field Displacement Diffusivity Test

B. R. Ragot
Heliophysics Research, P.O. Box 35611, Albuquerque, NM 87176, USA

Received 2021 June 30; revised 2021 September 12; accepted 2021 September 16; published 2022 March 15

Abstract

The estimate of the solar wind magnetic fields’ parallel correlation length, λ, be it from the measured fields’ correlation functions or their spectral power at “zero” frequency, have long pointed toward short values on the order of 0.01 au. Evaluation of the mean cross-field displacements (CFDs), however, fails to show the decorrelation and resulting diffusion at the expected scales, pointing instead toward λ values on the order of 0.1 au or more. In an effort to understand this “order-of-magnitude” discrepancy and reconcile the approaches using correlation functions and the CFD diffusivity test, both approaches are applied here, with renewed attention to the “details” as well as the broader sense of the calculations, to a large, 20 yr long set of magnetic field and flow data from the ACE spacecraft. It is found that solar wind intervals too short relative to λ are a likely reason for some underestimate through the correlation-function approach, causing a premature drop of the correlation functions. Once converged to their long-time limit, however, the correlation functions produce magnetic field correlation lengths very much consistent with the magnetic-field-line (MFL) correlation lengths of the diffusivity test, with nearly matching distributions of the correlation lengths corrected by the proper ratio of their theoretical estimates. The fields’ correlation lengths mostly range from 0.03 to 0.08 au, and the MFL correlation lengths from 0.04 to 0.3 au, with peaks at 0.075 and 0.15 au, likely due to nonlinear and quasilinear regimes of MFL wandering. As for the power-at-zero-frequency approach, it is doomed by the solar rotation.

Unified Astronomy Thesaurus concepts: Magnetic fields (994); Interplanetary turbulence (830); Space plasmas (1544); Heliosphere (711)

1. Introduction

The correlation length of the magnetic fields in the turbulent solar wind (SW) has traditionally been evaluated from the correlation function of the magnetic fields, be it a one-point time auto-correlation function1 or a multi-spacecraft two-point correlation function of the measured fields (e.g., Matthaeus et al. 1986; Richardson & Paularena 2001; Matthaeus et al. 2005, 2010; Weygand et al. 2011). An alternate yet mathematically related method evaluates the correlation length from the fields’ spectral power at “zero” frequency (Jokipii & Coleman 1968; Jokipii & Parker 1969), or somewhat equivalently, from the flattening frequency of the spectral power (Ragot 1999, 2006b, 2006c, 2018) and the resulting diffusion length of the magnetic field line (MFL) cross-field displacements (CFDs) or of the MFL spreading.

Both correlation-function and zero-frequency-power methods seem yield consistent results for the fields’ parallel correlation length in the SW, on the order of 10−2 au or 1–2 × 1011 cm (Matthaeus et al. 2005; Weygand et al. 2011; Jokipii & Coleman 1968; Jokipii & Parker 1969; see also Matthaeus et al. 1986; Matthaeus & Goldstein 1986; Matthaeus et al. 2010, who find a correlation length of ∼4–5 × 1011 cm, 2–4 times larger; Matsui et al. 2002, who find a correlation length of ∼0.8 × 1011 cm from a correlation analysis in the Fourier space, and Zhao et al. 2018, who estimate correlation lengths for slab and 2D magnetic fluctuations of ∼1011 cm and half of 1011 cm, respectively, through an Elsässer-variables approach). Evaluation of the MFL mean CFDs (Ragot 2006e), however, fails to show the expected decorrelation and resulting diffusion at such short scales of about 10−2 au and a few times 10−2 au, respectively. Instead, the extended supradiffusion of the mean CFDs past 1012 cm or 10−1 au in slow SW (and 3–5 × 10−2 au in fast SW where cross flow may precipitate the observed decorrelation; see Section 2.2.2 of Ragot 2006e) suggests correlations of the SW magnetic fields over scales much longer than anticipated. This issue needs to be resolved because the actual value of the magnetic fields’ correlation length is so fundamental to the proper modeling of solar energetic particles (SEPs). How SEPs and other SW energetic particles propagate through the SW magnetic fields very much depends on the actual value of the magnetic fields’ correlation length. Knowing the actual correlation length is also important to evaluate the predictive power, at Earth or at other locations of interest, of the SW observations made at L1 or elsewhere upstream in the SW flow (Crooker et al. 1982; Richardson et al. 1998; Richardson & Paularena 2001).

Parallel in “parallel correlation length” is sometimes defined along the direction of the main flow, the radial, sometimes along that of the main unperturbed magnetic field, the Parker spiral, and other times along that of the local mean magnetic field. While the definition of the parallel direction does have some bearing on the result for the parallel correlation length, variation in the definition of the parallel direction does not in itself produce such strong effects as to explain the apparent roughly order-of-magnitude discrepancy between estimates of the parallel correlation length. (The correlation length along the
parallel direction is usually the longest (Weygand et al. 2009), except in the fast SW where Dasso et al. (2005) and Weygand et al. (2011), who define the local mean magnetic field as the parallel direction, find it slightly (by a factor \(\sim 2\) at most) shorter than in the perpendicular direction.) In this paper, we use as parallel direction both the direction of the main flow, for the correlation-function analysis of Section 2, and the direction of the Parker spiral, when applying the CFD diffusivity-test method of Section 3 (see the beginning of Section 2.2). To compare the results of both methods, we translate the results for the correlation lengths from both methods back into times elapsed along the SW flow (see Section 4). We have to use the direction of the Parker spiral for the CFD analysis, because it is the background field direction in the study of SW MFL transport (be it through analytical calculations, data analysis, or numerical simulations). It is the most convenient choice from the theoretical viewpoint. Due to the solar rotation, the Parker spiral direction is the direction of the main unperturbed magnetic field, that is, the direction of the magnetic field in the SW flow were it perfectly regular (unperturbed or smooth).

If the variation in the definition of the parallel direction does not explain the apparent order-of-magnitude discrepancy between estimates of the parallel correlation length, what does? Here, we analyze twenty years of ACE magnetic field and flow data to investigate, “confirm,” and resolve the issue. By computing parallel “correlation lengths” from the same data set and through the same basic data analysis method, we confirm the potentially large discrepancy between methods based on correlation-function computation and the method based on the mean CFDs computation and diffusivity test. We propose a more cautious approach to the correlation-function computations/interpretation as well as to the application of the CFD diffusivity test. We find that both methods can actually be reconciled, with consistent results for the probability distribution functions (PDFs) of the computed correlation lengths, as well as for the individual correlation lengths. We start in Section 2 with the correlation-function method to evaluate a correlation length. A short general discussion of the method in Section 2.1 is followed in Section 2.2 with an application to the SW magnetic fields at 1 au using ACE magnetic field and flow data. To help distinguish fundamental scale-dependent behaviors of the correlation functions from the effects of SW transients and other large-scale inhomogeneities, we extend the scale-dependent study (both in separation scale and interval duration) to a couple of “pure” fast coronal-hole and slow SW intervals at solar minimum, in addition to the series of 180 37 day, 720 9 day, and over 7000 1 day intervals of the 1999–2018 ACE data set. We also present in Section 2.3 simple simulations and analytical estimates of the correlation functions, and using in situ projected power spectra of SW magnetic field fluctuations, we compare the results of simulation and theory to the correlation functions measured in a couple of SW intervals. Section 3 deals with the CFD diffusivity-test method and applies an improved method to the same ACE data set. It shows how the correlation length can be more systematically estimated and interpreted without recourse to spectral power computations, which are notoriously imprecise at low frequency. In Section 4, we reconcile both approaches and their correlation-length estimates. Our conclusion follows in Section 5.

2. Use of the Correlation Function to Evaluate a Correlation Length

2.1. Discussion of the Method

A physicist’s simple approach to the calculation of a correlation length through a correlation function \(C_{\lambda}(l) \equiv \langle A(z) A(z + l) \rangle / \langle A(z)^2 \rangle\) often consists in approximating the correlation function by an exponential, \(e^{-\lambda_\alpha l}\), and finding the length scale \(\lambda\) that best fits the correlations as function of the scale \(l\). Alternately, if the decrease of the correlation function \(C_{\lambda}(l)\) is badly fitted by an exponential, a correlation length can always be estimated as the length scale \(\lambda_{1/e}\) such that \(C_{\lambda}(\lambda_{1/e}) = 1/e\). A stricter definition involves the integral of the correlation function: \(\lambda_c \equiv \int_0^\infty dl \ C_{\lambda}(l)\), with often similar results (e.g., Matthaeus et al. 2005). The integral definition stands whether the correlation function \(C_{\lambda}(l)\) drops exponentially or not. For an exponential drop, \(\lambda = \lambda_c = \lambda_{1/e}\). If the drop of the correlation function is not exponential, the length scale \(\lambda_{1/e}\) may still be a worthy approximation. Due to lower statistics and higher uncertainties at the longer scales \(l\), it may sometimes be the only practicable estimate of a correlation length.

This is about all there is to say about the basics of the method (s) for the calculation of a correlation length. And we could very well proceed from here with the application to the ACE magnetic field data of Section 2.2. But we have also been wondering about how a statistician would evaluate a correlation length from the correlation function \(C_{\lambda}(l)\) and how the result would relate to the correlation length evaluated by a physicist. Would it give a different result? Could it help better understand the physics? And if it does not, can we put the question to rest? If our secondary questions about the statistician’s approach appear too involved or lacking interest to the reader, they can simply be skipped or filtered out. The main points of this paper are not affected by our parallel discussion of the statistician’s approach and calculation of \(L_{\alpha}\) in Section 2.2.4. From here, the reader only interested in the main points of this paper may skip to Section 2.2.

A statistician’s approach to the calculation of correlations through the use of the correlation function \(C_{\lambda}(l)\) is more concerned with the testing of the correlation (or noncorrelation) hypothesis on various length scales \(l\), that is, with the comparison of \(C_{\lambda}(l)\) to a critical value \(C_{\alpha,\alpha}\) of the correlation that depends on both the number \(N\) of observations used in the computation of \(C_{\lambda}(l)\) and the level \(\alpha\) of significance of the correlation, or confidence \((1 - \alpha) \times 100\%\) in the answer. Given a sample of \(N\) observations of the field \(A\) used in the computation of the correlation \(C_{\lambda}(l)\), finding a correlation greater than \(C_{0.01,N}\), the statistician will be able to conclude that the fields correlations are significant at the \(\alpha = 0.01\) level. The probability of the fields not being correlated on the length scale \(l\) is less than 0.01 (very low). In other terms, the statistician concludes with \((1-0.01) \times 100\% = 99\%\) confidence that the fields are correlated on the scale \(l\) (see a more detailed description below).

Now, if the statistician chooses to do the noncorrelation hypothesis testing with a value of the significance level \(\alpha' = 0.05\) or 0.1 instead of 0.01 and still finds that, for a new length scale \(l' > l\), \(C_{\lambda}(l')\) exceeds the critical value \(C_{\alpha',N}\), they will then conclude that the probability of the fields not being correlated on the new length scale \(l'\) is less than 0.05 or 0.1, or equivalently, conclude with 95 or 90% confidence that the fields are correlated on the scale \(l'\). At what point, from the
The next two paragraphs are our attempt at summarizing the practical physicist’s point of view, will the statistician be able to conclude that the fields are not correlated? At a significance level of 0.25, 0.5, 0.75? Could this be a better approach to determining the decorrelation length? Is it practicable? What does it tell us about the “physicist’s approach”?

By computing large statistics of the correlation functions from the 20 yr long ACE data set, we will see that the earlier results (e.g., Matthes et al. 2005; Weygand et al. 2011) based on the computation of correlation functions are consistent with the simple physicist’s approach with an estimate of \( \lambda_{1/e} \) from \( C_B(\lambda_{1/e}) = 1/e \), which seemingly gives a relatively short estimate of the correlation length, an estimate for which the statistician would conclude with better than 99% confidence that the fields are still correlated, or find that the probability of the fields not being correlated on that scale is much less than 0.01. This is not necessarily an issue if one understands the correlation length to be the scale over which the fields are still strongly correlated but beyond which (or a few of which) the correlations sharply drop.

Coming back to the basics, a correlation coefficient \( r \) cannot be used directly to indicate the degree of correlation (e.g., Bevington 1969). This would seem to disqualify the “simple physicist’s approach” from the start, but it does not. The integration of the correlation function over all scales to estimate the correlation length \( \lambda \) is not affected by this limitation, because it does not make claims about the actual degree of correlation at any given scale. It estimates a scale over which strong correlations exist, and beyond which strong correlations disappear. The estimate \( \lambda_{1/e} \) of the correlation length based on the exponential approximation would be baseless on its own, but as an approximation of the integral result, has merit, provided the correlation function is indeed not too far from an exponential. Although the approach based on the integral definition of the correlation length is in principle always valid, it is sometimes difficult to apply because of slow convergence and/or large error bars in limited statistics (at the larger scales). The uncertainty in the correlation function at the larger scales may, in some cases, make the \( \lambda_{1/e} \) estimate the only practicable estimate. So some kind of “backup” method for a consistency check would be welcome. This is also why we are tentatively exploring the possibilities offered by a “statistician’s approach,” and explaining the rudiments of that approach below.

In a common test of the correlation coefficient \( r \) known as the zero-correlation hypothesis testing, the value of \( r \) derived from a sample of \( N \) data points or fields measurements is interpreted in relation to the probability distribution (Pugh & Winslow 1966; Bevington 1969)

\[
P_c(r, N) = \frac{\Gamma[(\nu+1)/2]}{\sqrt{\pi} \Gamma(\nu/2)} (1 - r^2)^{\nu-2}/2
\]  

(1)

that any random sample of uncorrelated data points would yield a linear-correlation coefficient equal to \( r \) from a parent population that is completely uncorrelated, i.e., with zero correlation, and to its integral

\[
P_c(r, N) = \int_{|r|}^1 d\rho \, P_c(\rho, N - 2),
\]  

(2)

which effectively gives the probability that a random sample of uncorrelated data points would yield a correlation coefficient as large as or larger than \( r \), for a one-tail (positive correlation) test. In Equation (1), \( \nu = N - 2 \) is the number of degrees of freedom. A small value of \( P_c(r, N) \) indicates that the observed data are probably correlated.

In practice, a value \( \alpha \) called a significance level is pre-selected for the integral \( P_c(r, N) \) of Equation (2) and the corresponding critical value \( r_\alpha \) for which \( P_c(r_\alpha, N) = \alpha \) is estimated \( (r_\alpha \) is the \( C_{\alpha N} \) of our earlier discussion in this section), to which the computed correlation coefficient \( r \) is then compared. Such a comparison indicates whether or not it is likely, at an \( \alpha \) significance level, that the data points could represent a sample derived from an uncorrelated parent population. If \( r \geq r_\alpha \), the zero-correlation hypothesis testing at an \( \alpha \) significance level fails; it is concluded that the observed fields are correlated at an \( \alpha \) significance level. Figures 1 and 2 show in linear and logarithmic scale, respectively, the critical values \( r_\alpha \) of the correlation coefficient, as functions of the number \( N \) of data points between 3 and \( 3 \times 10^3 \), for a series of significance levels \( \alpha \) between 0.001 and 0.4.

2.2. Application to the ACE Magnetic Field Data

Here and in the following section we analyze 20 yr (1999–2018) of 1 s resolution magnetic field data from the MFI magnetometer (Smith et al. 1998) on the ACE spacecraft.
We also make use of the 64 s resolution SW flow data from the ACE SWEPAM instrument (McComas et al. 1998). To reduce memory use and speed up computations, we resampled the magnetic field data to a 3 s resolution. The ACE data were downloaded in RTN coordinates (in which \( \mathbf{R} \) is along the radial direction, from the Sun to the spacecraft, pointing away from the Sun, \( \mathbf{T} \) is the cross product of the Sun’s rotation vector and \( \mathbf{R} \), and \( N \equiv \mathbf{R} \times \mathbf{T} \) completes the triad). They were converted to \( \chi \rho \zeta \) coordinates, in which \( z \) is along the local Parker spiral, pointing away from the Sun, \( x \) is the cross product of the Sun’s rotation vector and \( z \), and \( y \equiv z \times x \) is identical to \( N \).

Because we are using data from only one spacecraft (ACE), here we compute auto-correlation functions \( C_{B_{\chi}}(l) \), defined for a field \( B_{\chi} \) as

\[
C_{B_{\chi}}(l) \equiv \frac{\int_0^{L-1} dz [B_{\chi}(z) - \bar{B}_{\chi}][B_{\chi}(z + l) - \bar{B}_{\chi}]}{\int_0^L dz [B_{\chi}(z) - \bar{B}_{\chi}]^2},
\]

where \( \bar{B}_{\chi} \) is the field’s mean value on the interval of length \( L \) (see also footnote 1). In the following, whenever we refer to these functions, “auto” is omitted.

The measurements at one single spacecraft at 1 au allow one to estimate the fields’ correlations most naturally along the nearly radial line of the SW flow past the spacecraft. Here in Section 2.2, we measure the fields’ correlations and correlation time or length along the SW flow; parallel correlation refers to the correlations measured along SW flow lines, and parallel correlation length to the correlation time along these SW flow lines multiplied by the SW speed. In Section 3, we will estimate the correlation lengths along the Parker spiral with a simple translation of the time intervals in the SW flow past the spacecraft into local Parker spiral lengths, \( \Delta z = V_{SW} \Delta t / \cos \psi \), where \( \psi \) is the angle between the radial and Parker spiral directions. To test the consistency of the two in Section 4, we will translate both results for the correlation lengths back into times along the flow. Again, in this Section 2.2, parallel correlation naturally refers to the correlations measured along SW flow lines, and parallel correlation length to the correlation length along these SW flow lines.

### 2.2.1. Duration \( T \) of the Intervals of Analysis

When computing the correlation at a scale \( l \) of a data sample of duration \( T \) and length \( L = V_{SW}T \), perhaps the first question that arises concerns the choice of \( T \). Clearly, demonstrating correlations on a scale \( l \gg L \) would be challenging, so \( T \) has to be chosen to be long enough. Earlier studies based on the computation of the correlations (see references in the Introduction) have, as far as we know, used \( T \approx 1 \) day or a fraction thereof. An interval duration \( T \approx 1 \) day roughly corresponds to a length \( L \approx 3 \times 10^{12} \) cm or 0.2 au. Would this length be sufficient to demonstrate correlations on a scale on the order of 0.1 au? How would too short a length affect the shape of the correlation function? As a test, here we compute the correlations at a series of scales \( l \) on intervals of three different durations, \( T \geq 1 \) day, using the \( B_{\chi} \) component of the 3 s resolution ACE magnetic field data set. Figure 3 shows for the years 1999 (black), 2005 (green), 2011 (blue), and 2017 (red) the yearly averages of the correlation functions computed on series of intervals of duration \( T \approx 37 \) days (dotted lines), \( T \approx 9 \) days (solid lines), and \( T \approx 1 \) day (dashed lines). For clarity of Figure 3, the lines for year yr are shifted down by \( 0.05 \times (yr - 1999) \). The horizontal lines indicate the zero line for the correlations of the year with matching color. The \( l \) values in Figure 3 are estimated with the year-averaged \( V_{SW} \).

The main conclusion to be drawn from Figure 3 is that analyzing the data on series of time intervals that are too short causes the correlations to drop prematurely. Whereas increasing \( T \) past 9 days has a relatively minor effect at higher \( l \), the underestimate of the correlations computed from 1 day intervals can be striking beyond a fraction of \( 10^{11} \) cm.

This could be a major reason for the low estimates of the correlation length by previous authors. However, the length of the longer intervals (9 and 37 days) raises the possibility that larger-scale inhomogeneities, corotating interaction regions (CIRs), coronal mass ejections (CMEs), and other solar transient events could be playing a role in increasing the correlations on the longer scales. Here in Figure 4, we extend the analysis to a comparison of the correlation functions computed, for the year 2009 of extreme solar minimum, with a series of eight \( T \) durations between 1.75 and 224 hr (with successive \( T \) values in a ratio of two) plus \( T = 896 \) hr or 37.3 days. Figure 4 clearly confirms the dependence of the correlations on the duration \( T \) of the intervals of analysis, the
dependence being strong up to a couple of days and then dampening, as already observed past 9 days in Figure 3.

Our choice of the year 2009 of extreme solar minimum, with practically no CMEs or solar transient events, makes these events an unlikely explanation for the observed $T$-dependence. Still, we cannot preclude at this point that transitions from slow to fast wind may play a role. To preclude CIRs as well as CMEs and other solar transient events, we also considered a series of well-defined fast coronal-hole and slow SW intervals. A few can be found in the years 2008 and 2009 of low solar minimum that are at least 112 hr or 4.7 days long.

Figures 5 and 6 show for series of $T$ durations between 1.75 and 112 hr, the correlations versus scale $\tau$ in the fast and slow SW observed between days 15 and 19.7 of year 2008 and days 90 and 94.7 of year 2009, respectively, well away from stream boundaries. They are representative of the other clearly defined fast and slow SW streams observed in the years 2008 and 2009, though there are variations between streams. To illustrate that variability, Figure 7 shows the correlations versus scale $\tau$ for a series of $T$ durations on another interval of slow SW, observed between days 10.5 and 15 of 2009. The $l$- or $\tau$-dependence of the correlations in the fast SW stream (Figure 5) is most definitely not exponential, and greatly differs from that in the slow SW streams (Figures 6 and 7). Most importantly, the $T$-dependence is again clearly confirmed, and at this point, should be considered a very likely basic feature of SW magnetic field correlation functions from the observational point of view. Section 2.3 further investigates the issues of the $T$-dependence and of the (nonexponential) $l$- or $\tau$-dependence with simple numerical simulations and from the theoretical point of view, with simple analytical estimates.

### 2.2.2. Sampling Rate and Correlations

When computing the correlation at scale $l$ of a data sample of duration $T$ and length $L = V_{SW}T$, the question also arises as to how the time resolution of the data will affect the result of the correlation computation. Obviously, a higher resolution or higher sampling rate will give a higher number $N$ of data points on the total length $L$ of the interval; and we have seen that the statistician’s interpretation of the correlation depends on the number $N$. But when we choose a scale $l$ to compute the correlation $r = C_{R}(l)$, does the result for $r$ change with the number of intermediate points (between points separated by a distance $l$) included in the sample? If the data are well correlated on that scale, it should not. We check from the ACE data set that it does not.

Figure 8 shows the results of correlation functions $C_{R}(l)$ computations made by sampling the magnetic field data at a $\delta t \approx 3$ s resolution (dotted lines), whatever the value of the
length scale $l$, and at a scale-dependent resolution $\Delta t \approx l/V_{SW}$ (solid lines), where the $V_{SW}$ used is the $T$-averaged SW flow speed. The correlation functions are computed on intervals of duration $T \sim 9$ days. The results shown in Figure 8 are the averages of the 36 successive intervals computed for each year running from 1999 to 2018. The $l$ values in Figure 8 are estimated with the year-averaged $V_{SW}$. For clarity of Figure 8, the lines for year yr are shifted down by 0.05 × (yr − 1999). The two horizontal lines indicate the zero line for the first and the last year (black and red, respectively). The correlations $C_B(l)$ have been computed at series of 18 values of $l$, shown by the + symbols, each separated by a factor 2 (only the first six appear on Figure 8). The individual points are joined for clarity of the figure. We see that the solid and dotted lines practically superpose until the noise in the lower-resolution solid line increases dramatically, beyond about $10^{12}$ cm.

The results for individual 9 day long intervals show a similar match, up to a fraction of $10^{12}$ cm. There is a lot of shape variation between individual intervals of SW, though, as illustrated by Figure 9. Figure 9 shows the correlations computed for each of the 36 successive 9 day long intervals of year 1999, the average of which is displayed as a black solid (dotted) line in Figures 3 and 8. In Figure 9, $l$ is estimated with the $T$-averaged $V_{SW}$. We see that there is quite a spread in the value of $C_B(l)$ over time. For clarity of Figure 9, only the less noisy lines computed from the 3 s resolution data are shown.

### 2.2.3. Estimating the Correlation Lengths

With $T \geq 9$ days, the year-averaged correlations $C_B(l)$ computed from the 3 s resolution data are smooth enough down to $C_B(l) \approx 0.2$ or even 0.1 for a study of the correlations at the scales $l \leq 10^{12}$ cm. Without year averaging, most ($T = 9$ days) correlations remain smooth enough down to somewhere between $C_B(l) = 0.2$ and 0.4, at scales $l \lesssim 10^{11}$ cm to a few times $10^{11}$ cm. To compute the correlation length from its integral definition, $\lambda_c = \int_0^\infty dl \ C_B(l)$, the higher statistics of the longer intervals and higher resolution are really needed to narrow the error bars at the higher $l$ (or lower $C_B(l)$).

If the drop of the correlation functions $C_B(l)$ were exponential, there would be no need to worry about the uncertainties at higher $l$. One could just fit an exponential to $C_B(l)$ from low $l$. The magnetic fluctuations observed in the SW are not monoscale, however, and even though, e.g.,

Wicks et al. (2010) concluded otherwise, the correlation functions $C_B(l)$ do not drop exponentially. The extended power laws in the measured power spectra of magnetic fluctuations are the main reason/argument for this nonexponential drop (see the relation between power spectra and correlation functions in Section 2.3; see also Figure 4 of Ragot 2006c for an illustration of the self-similarity of a field with a power-law spectrum). On broader ranges of scales, exponential fits can be poor, particularly in the fast SW, which has relatively hard power spectra of magnetic fluctuations (see Figures 5–7, and 10–11 with higher statistics; see also Section 2.3). Until the actual correlation length is reached, the decrease of $C_B(l)$ is closer to one minus a slow power law, with the main contribution to the integral possibly coming from the highest values of $l \leq \lambda_c$. Therefore, the main difficulty in estimating $\lambda_c$ from the correlation-function integral is to estimate the correlations at long enough $l$ with a low enough uncertainty. When computing the $\lambda_c$ and their distributions from the correlation-function integral, we must use the higher ($\delta t \approx 3$ s) resolution data to minimize the uncertainty. We must also include long enough $l$ scales in our correlation analysis, that is, choose $T$ long enough.
In the statistician’s approach, the match of the yearly averages of the correlations in Figure 8, as well as of the correlations on individual intervals, justifies that we sample the data at the better resolution (3 s) to compute the correlation \( r = C_{B_\lambda}(l) \) with a higher precision, but use the number \( N = L/l \) of data points to interpret the correlation. Using the greater \( N \) of a higher sampling rate would result in lower critical correlations \( r_c \) (see Section 2.1). This would result in longer correlation-length estimates because at a given \( l \), the conclusion of a negative no-correlation test would more easily be reached. By using the \( N \) values of the lowest possible sampling rate \((V_{SW}/l)\), we are ensuring that decorrelation is detected as soon as it occurs, at the lowest possible \( l \). We compute the correlations with the better resolution, because good precision is absolutely needed at longer \( l \) for higher values of the significance level \( \alpha \). We show in Figures 10 and 11 the results of computations made with the 3 s resolution data, but in the statistician’s approach, interpret the results with \( N = L/l \), as if the resolution were \( l \).

If we sampled the data at the scale \( l \) from a higher-resolution data set, it would seem natural to also average our data on the scale \( l \) before sampling the data at that scale. So we do that too, and compare the results. Figures 10 and 11 show the correlations \( r = C_{B_\lambda}(l) \) computed with and without pre-averaging on the scale \( l \), for \( T \sim 9 \) and 37 days, respectively. Here, we find significant differences between the results with and without pre-averaging. We do not make any firm conclusion here as to which computation leads to the best estimate of the correlations (this will be addressed later in Section 4), but it is clear that the computation with pre-averaging on the scale \( l \) would give the longest estimates for the “correlation length.” The pre-averaging smooths out the field, making it more likely to measure higher correlation.

In addition to the correlations computed from the 3 s averaged (solid lines with crosses\(^3\)) and the l-averaged (dashed lines with crosses) magnetic field data, Figures 10 and 11 also show as dotted lines the exponentials that reach \( 1/e \) at the same \( l = \lambda_{1/e} \) as the 3 s averaged field correlations. From these figures it is clear that exponentials are very poor fits to the correlations. Still, the intersection points \((\lambda_{1/e}, 1/e)\) provide a common and easy estimate of correlation length, with the clear advantage of not being affected by uncertainties at the highest scales \( l \). We note also that the main departure from the exponential passing through \((\lambda_{1/e}, 1/e)\) is often at the shorter scales \( l \) (see also Figure 5), meaning that the exponential scale \( \lambda_{1/e} \) is actually closer to the longer relevant scales than an attempt to fit \( C_{B_\lambda}(l) \) at shorter \( l \) would yield.

Figure 12 gives in dotted lines the distributions of the lengths \( \lambda_{1/e} \) for each of the interval durations \( T = 37 \) days (top panel), \( T = 9 \) days (middle panel), and \( T = 1 \) day (bottom panel), both for the 3 s averaged fields (black dotted histograms) and l-averaged fields (red dotted histograms). The broad width of these distributions reflects the high variability of the correlation functions between individual intervals (see Figure 9). Here and throughout this paper, the histograms are computed from the complete series of intervals of duration \( T \) over the 20 years of our data set, not from the yearly averaged results. The distributions’ widths may also be affected by a solar-cycle effect (see end of Section 4).

---

\(^3\) The crosses indicate where the correlations were actually computed.

---

Figure 12 also shows in solid lines the histograms of the correlation lengths \( \lambda_c \) computed from their integral definition, again for each of the interval durations \( T = 37 \) days (top panel), \( T = 9 \) days (middle panel), and \( T = 1 \) day (bottom panel), and both for the 3 s averaged fields (black histograms) and l-averaged fields (red histograms). Both \( \lambda_c \) and \( \lambda_{1/e} \) distributions are roughly consistent, within a factor of 2. (This near consistency is likely due to the fact that much of the nonexponential behavior of the correlation functions occurs before they reach \( 1/e \).) The lengths \( \lambda_{1/e} \) appear to often underestimate the correlation lengths \( \lambda_c \), by a factor on the order of 2 on average for \( T = 37 \) days, and by a smaller factor \((\approx 1.5)\) for \( T = 9 \) days. For \( T = 1 \) day, the distributions are nearly identical. Also, the distributions shift down as the duration \( T \) of the SW intervals decreases, at a rate that appears to accelerate downward.

---

2.2.4. Length Scales \( L_\alpha \) for Correlation at Significance Level \( \alpha \)

The rising lines in Figures 10 and 11 give the critical values \( r_c \) (see Section 2.1) of the correlation coefficient as functions of the length scale \( l \), for significance levels \( \alpha = 0.01, 0.05, 0.10, 0.25, \) and \( 0.40 \), as indicated to the right of each \( r_c \) line. Where the correlation functions intersect the \( r_c \) lines, gives the length scales \( L_\alpha \) at which we know the data to be correlated at a significance level \( \alpha \). Figure 13 shows the histograms of these length scales \( L_\alpha \). The increasing uncertainties with larger \( l \) and lower \( r \) in the estimates of \( r \) (on individual intervals) cause the distribution of \( L_\alpha \) to broaden dramatically past \( \alpha = 0.25 \) (especially for the l-averaged data, for which correlations must be computed with much lower statistics), so the result for \( L_{0.4} \) is not shown in Figure 13.

The four panels of Figure 13 show the \( L_\alpha \) histograms for the 3 s averaged fields (top panels) and the l-averaged fields.
correlated at a significance level $\alpha$. Top panels: computed with the 3 s resolution data. Bottom panels: computed with the scale-l-averaged data. Left panels: over about 740 9 day long intervals. Right panels: over about 180 37 day long intervals. Black solid line histograms: for $\alpha = 0.01$. Green dashed line histograms: for $\alpha = 0.05$. Red solid line histograms: for $\alpha = 0.10$. Blue dashed line histograms: for $\alpha = 0.25$.

(bottom panels), for $T \sim 9$ days (left panels) and for $T \sim 37$ days (right panels). The statistics for the computation of the correlations are higher at 37 days, but accurate enough correlations must be estimated down to lower values of the correlations to find the intersection points with the $r_\alpha$ lines (compare Figures 10 and 11, keeping in mind that the correlations shown in these figures are yearly averages. The results for individual intervals are much noisier at higher $t$—see Figure 9 for $T = 9$ days). So the results are not necessarily more accurate for $T \sim 37$ days. Still, aside from the red and blue distributions of the lower right panel, which show significant fluctuations (presumably due to the uncertainties in determining the intersection points, but perhaps not—see below), all distributions in Figure 13 appear to be well defined, with only one major peak. We note that the two histograms (left bottom and top panels) for $\alpha = 0.01$ and $T = 9$ days bear a strong resemblance to the black and red histograms of Figure 12’s middle panel. This we believe adds credence to the validity of the correlation lengths estimated through the correlation functions (provided that the duration $T$ of the data intervals is long enough (> a few days)).

The resemblance is also strong for $T = 37$ days, with a shift of about 40%. However, it is the red dotted line histogram of the top panel of Figure 12 that best matches the $L_{0.01}$ histogram of the bottom-right panel of Figure 13. (The red solid line histogram fails to drop by a few times 0.1 au.) On the one hand, this could be an example where $\lambda_{1/e}$ yields a more reliable estimate of the correlation length than the integral of $\lambda_c$, due to high uncertainties at the longer scales $l$. On the other hand, the tail in the red solid line distribution could be due to real correlations (see Figure 27 in Section 3.6, and Figures 32 and 33 in Section 4, which show a better match—weaker slope—for $L_c$ than for $\lambda_{1/e}$). If these correlations are due to markedly nonexponential behavior at longer $l$ (as opposed to a global shift to higher $l$ of the entire correlation function) for a subset of intervals, neither estimates of $\lambda_{1/e}$ nor of $L_{0.01}$ would be able to reflect the existence of such correlations because $\lambda_{1/e}$ and $L_{0.01}$ are obtained as the scales $l$ where the correlations reach finite values $1/e$ and $r_\alpha(0.01, L/l)$, respectively. Neither distributions of $\lambda_{1/e}$ nor of $L_{0.01}$ should be expected to have a tail. But the series of $L_\alpha$ distributions for higher $\alpha$ (decreasing $r_\alpha(0.01, L/l)$) may be able to reflect a nonexponential drop of the correlations at the longer scales $l$ (extended correlations) for a subset of intervals, and to confirm the existence of a tail. They actually seem to (see red and blue histograms in the bottom-right panel of Figure 13, also extending past 1 au).

Again (see Section 2.1), there is no issue in the fields being correlated at a 0.01 significance level on the correlation length $\lambda_c$ or even up to 1.5 $\lambda_c$. The correlation length is the length scale over which the fields are strongly correlated but beyond a few of which the correlations drop. It should take several correlation lengths for the correlations to be reduced to a “negligible level.” Even though the low statistics do not allow...
us to verify with much confidence, from the individual data intervals, on what scale the correlation drop actually occurs, the year-averaged results indicate a drop to $\alpha = 0.4$ by about 6 correlation lengths (see footnote 4). So we believe that the “statistician’s approach” does provide a useful “backup” method for a consistency check. Its results add credence to the results obtained by the “physicist’s approach.” It also clarifies the meaning of correlation length. Beyond correlation-length estimates, the length scales $L_n$ might also find some use in dealing with questions of fields’ predictability or the predictive power of SW upstream measurements.

2.3. Simulations and Analytical Estimates of SW Magnetic Field Correlations

Here, we further investigate through simple numerical simulations and analytical estimates the dependences of the correlation functions on the separation scale $\tau$ and the duration $T$ of the intervals of analysis. For clarity of the presentation, in this section, we consider the correlations as functions of the timescale $\tau$, since the data are direct functions of time and the length scale $l$ is only deduced from $\tau$.

For a data interval of duration $T$, the correlation function,

$$C_{\Lambda}(\tau) \equiv \frac{1}{T} \int_0^T dt \left[ A(t) - \bar{A} \right] [A(t + \tau) - \bar{A}],$$

(4)

at the timescale $\tau$, of the field $A$, is related to the projected (measured) power spectrum $P_{\Lambda}(\nu)$ through (see the Appendix)

$$\sigma(\tau) C_{\Lambda}(\tau) = \frac{\int d\nu e^{i 2 \pi \nu \tau} P_{\Lambda}(\nu)}{\int d\nu P_{\Lambda}(\nu)},$$

(5)

where $\sigma$ is a window function of width $T$. This results from the very definition of the power spectrum as the Fourier transform of the correlation function (e.g., Jokipii & Parker 1969).

Assuming a piecewise power-law spectrum with, on each frequency interval $[\nu_j, \nu_{j+1}]$,

$$P_{\Lambda}(\nu) = (\eta/\nu_j)^{\nu_j} P_{\Lambda}(\nu_j),$$

(6)

and a flat spectrum below $\nu_j$, the right-hand side of Equation (5) can be integrated to give, if $T$ is very much longer than $\nu_j^{-1}$,

$$\sigma(\tau) C_{\Lambda}(\tau) \int_0^{+\infty} d\nu P_{\Lambda}(\nu) = \nu_0 P_{\Lambda}(\nu_0) \frac{\sin(\pi \nu_0 \tau)}{\pi \nu_0 \tau} \cos(\pi \nu_0 \tau) + \text{Re} \sum_{j \geq 0} \frac{\nu_j P_{\Lambda}(\nu_j)}{(2\pi \nu_j)^{\nu_j + 1}} [H_j(\nu_j, \tau) - H_j(\nu_{j+1}, \tau)],$$

(7)

where

$$H_j(\nu, \tau) \equiv e^{i (a_j + 1) \tau^2} \Gamma(a_j + 1, -i2\pi\nu\tau).$$

(8)

If $T$ is not much longer than $\nu_j^{-1}$, a heuristic way of accounting for the finite duration of the data interval and the resulting finite lower frequency of the projected spectrum is to introduce a lower cutoff frequency $\nu_T = (\eta T)^{-1}$ in the integral over $\nu$, leading to

$$\sigma(\tau) C_{\Lambda}(\tau) \int_0^{+\infty} d\nu P_{\Lambda}(\nu) = \nu_0 P_{\Lambda}(\nu_0) \frac{\sin(\pi \nu_0 \tau)}{\pi \nu_0 \tau} \cos(\pi \nu_0 \tau)$$

$$\times \text{Re} \sum_{j \geq 0} \frac{\nu_j P_{\Lambda}(\nu_j)}{(2\pi \nu_j)^{\nu_j + 1}} [H_j(\nu_j, \tau) - H_j(\nu_{j+1}, \tau)],$$

(9)

where $H$ is the Heaviside or step function. By comparing the results of numerical simulations and theoretical estimates, we find that $\eta = 2^{1/2}$ gives a reasonable match.

In order to verify the validity of these theoretical expressions, we now simulate the field $A(t)$ from its projected power spectrum $P_{\Lambda}(\nu)$. Any geometry could be assumed for the distribution of wavevector directions, provided that the projection of the 3D power onto the measurement or flow direction produces the projected (measured) spectrum $P_{\Lambda}(\nu)$. But the easiest and computationally most economical geometry is the slab geometry, wherein all wavevectors are aligned in one direction, that of the flow. So we model the field $A$ as

$$A(t) = \nu_m^{1/2} \sum_n P_A^{1/2}(\nu_m) \cos(2\pi \nu_m t + \alpha_n),$$

(10)

with random phases, $\alpha_n$, and the discretization frequency,

$$\nu_m = (10 T_{\text{simulation}})^{-1} \ll (\nu_0, T^{-1}),$$

(11)

for a simulated field of total duration $T_{\text{simulation}} = 40$ days, and intervals of data correlation analysis of duration $T$ ($<T_{\text{simulation}}$). The conditions of Equation (11) guarantee that there are no periodicity effects in the simulated field.

Using the above model, we compute the field $A$ at a series of times $t_p = p \delta t$, with the same resolution $\delta t$ of 3 s as for the SW data. To improve the statistics, we compute one hundred independent realizations of the field $A$, with uncorrelated series of random phases $\alpha_n$. We then apply the same method of correlation analysis as for the SW in situ field, choosing a series of interval durations $T$ between 1.75 and 224 hr, with successive $T$ values in a ratio of two, plus 896 hr. In the example of Figure 14, we use for the projected spectrum the green spectrum (with $a = -0.9999$ instead of $-1$, to allow computation of the incomplete Gamma functions of Equation (8)) that will be calculated in Section 3.7 for the third 37 day interval of the year 2009 (see Figure 28 in Section 3.7).

Figure 14 shows the results of our correlation analysis for the simulated field (dotted lines) and their theoretical, analytical approximations (dashed lines) with a lower cutoff frequency $\nu_T = (2^{1/2} T)^{-1}$ to account for the finite duration $T$ of the data intervals in the correlation analysis. The thin black solid line is computed from Equations (7–8) without a lower $\nu_T$ cutoff, and the thick solid line shows the approximation wherein an upper cutoff frequency,

$$\nu_T = (2^{1/2} 2\pi T)^{-1},$$

(12)
is introduced in the integral \( \int d\nu e^{i2\pi \nu \tau} P_\text{A}(\nu) \) of Equation (5), which is estimated as
\[
\int d\nu e^{i2\pi \nu \tau} P_\text{A}(\nu) \approx \int_0^{\nu_{\text{max}}} d\nu P_\text{A}(\nu). \tag{13}
\]

The correlation functions are clearly not exponential, and they strongly depend on the duration \( T \) of the data intervals.

We also compare in Figure 15 the results of our simulations with those of our in situ data analysis. The match between in situ data and simulation results is quite good above separation scales \( \tau \) of a few times \( 10^{-2} \) hr. In both figures, the upper frequency \( \nu_{\text{max}} \) in the simulations and analytical estimates is \( 4 \times 10^{-3} \) Hz, which, by producing a smoother field, causes the correlations to remain close to one over an extended range of timescales \( \tau \) (up to \( \tau \sim 10^{-2} \) hr). We checked that more realistic simulations with a more extended inertial range produce a more gradual increase of the correlations toward the lower timescales \( \tau \), with a close match at short separation scales \( \tau \), but they are much slower and therefore more difficult to iterate a hundred times (over the entire range of timescales \( \tau \)) for good statistics.

In Figure 15, the in situ data and averaged simulation results show significant differences at longer \( T \) (\( \gg 1 \) day) and \( \tau \). We find that this is due to the lower statistics of the in situ data analysis at longer \( T \), and the strong variability of the correlations with the field realization. The thin dashed–dotted lines show one example of a field realization in our simulation of one hundred such fields that gives a fairly good match to the in situ data analysis results at all \( T \). (We did not look for the best matching realization; the first one, out of one hundred, was just good enough.)

We also compare in Figures 16 and 17 the results of our simulations with those of our in situ data analysis for the fast and slow SW intervals of Figures 5 and 6. For these shorter intervals, we used \( \nu_{\text{max}} = 10^{-1} \) Hz and fewer fields’ realizations (40 instead of 100). The match is reasonably good, but (unlike in Figure 15) did not benefit from knowledge gained in our study of Section 3 (in particular Section 3.7) about the low-frequency power spectrum.

From Figures 14–17, as well as Figures 3–7 in Section 2.2, it is clear that the correlations do very much depend on the
duration $T$ of the intervals of analysis, and a variation in the type of turbulence or the appearance of structures such as CMEs in the longer data intervals may not be blamed for all of this $T$-dependence. It is just the nature of extended piecewise power-law spectra to produce such a dependence on $T$.

3. Use of the Cross-field Displacements to Evaluate a Correlation Length

3.1. CFDs and Diffusivity

The amount of correlation, or lack thereof, indicative of actual decorrelation of the fields is not known, as hinted by our discussion of Section 2 (Sections 2.1 and 2.2.4 in particular). But a straightforward way to assess whether decorrelation has occurred is to test for the diffusivity of the field lines. When decorrelation occurs on a scale $L_c$, by definition of decorrelation, the successive field-line cross-field displacements (CFDs) $(\delta r)_n$ on the background field aligned scales $(\delta z)_n = l > L_c$ are independent. As a result, the spreading over $N = \Delta z/l = \sum_{n=0}^{N} (\delta z)_n/l$ successive steps,\(^5\)

$$\langle (\Delta r)^2 \rangle = \left\{ \sum_{n=0}^{N} (\delta r)_n^2 \right\}$$  \hspace{1cm} (14)

reduces to

$$\langle (\Delta r)^2 \rangle \approx \sum_{n=0}^{N} \langle (\delta r)_n^2 \rangle \approx N \langle (\delta r)^2 \rangle \approx \frac{\Delta z}{l} \langle (\delta r)^2 \rangle,$$  \hspace{1cm} (15)

where the correlations $\langle (\delta r)_n \langle (\delta r)_n \rangle$ between different steps have been neglected. It increases linearly with the distance $\Delta z$ elapsed along the background field direction. It is said to be diffusive. The MFL cross-field diffusion starts at a few correlation lengths of the fields (see further below, around Equation (32), for a refining of “a few”). On the shorter scales $\Delta z_c$, the fields’ correlations cause supradiffusion of the MFLs, $\langle (\Delta r)^2 \rangle \propto (\Delta z)^{\alpha} \text{ with } \alpha > 1$ (Ragot 1999, 2001, 2006c, 2006d, 2018).

So, an alternate (substitute) approach for evaluating the fields’ correlation length uses the MFL CFDs instead of the fields’ correlation function. It consists in determining the length beyond which the MFLs become diffusive, that is, have a spreading $\langle (\Delta r)^2 \rangle$ that increases linearly with $\Delta z$. Here, we use the ACE data set to test for the MFL diffusivity on a broad range of length scales and determine the correlation length $L_c$. But some more theoretical background will guide us in our data analysis, by helping us determine which quantity to consider to best determine $L_c$.

3.2. Some Theoretical Background

In the quasilinear calculation for the MFL wandering across a background field $B_0$, the magnetic field perturbation $\delta B$ is assumed small enough that $\Delta r$, the MFL perpendicular displacement across $B_0$ or CFD can be neglected in the turbulent field (Jokipii & Parker 1968, 1969). The transverse turbulent field is evaluated along a line that follows the background field direction instead of a real MFL. This approximation is known as the quasi-linear approximation. The same quasi-linear approximation is made in the generalized quasilinear (GQL) calculations that extend the quasilinear calculations to all field aligned length scales $\Delta z$ (Ragot 1999, 2001, 2006c, 2006d, 2018). Unlike in the original quasilinear calculation, however, the GQL calculations do not assume a length scale $\Delta z$ much longer than a parallel correlation length $L_c$. The GQL calculations establish that $L_c = k_0^{-1}$, where $k_0$ is the parallel wavenumber below which the projected turbulence spectrum $P_\perp(k_0)$ is flat. In the quasilinear or GQL approximation, diffusion occurs beyond $k_0^{-1}$.

In the quasilinear regime of turbulence, where the quasilinear approximation is justified (by definition), the mean square CFD,

$$\langle (\Delta r)^2 \rangle \equiv \langle (\Delta x)^2 \rangle + \langle (\Delta y)^2 \rangle,$$  \hspace{1cm} (16)

is shown to be the triple integral (see Equation (7) in Ragot 1999 or Ragot 2006c),

$$\langle (\Delta r)^2 \rangle = 2 \int_{z_0}^{\Delta z} dz' \int_{z_0}^{\Delta z} dz'' \times \int_{0}^{k_\alpha} dk_\parallel \cos [k_\parallel (z' - z'')] P_\perp(k_\parallel),$$  \hspace{1cm} (17)

of the background field projected power spectra,

$$P_\perp(k_\parallel) \equiv k_\parallel^3 \int d\mathbf{k}_\perp (\vec{b}_x^2 + \vec{b}_y^2)(k_\parallel, \mathbf{k}_\perp),$$  \hspace{1cm} (18)

which easily integrates to

$$\langle (\Delta r)^2 \rangle = 4 \int_{0}^{k_\alpha} dk_\parallel k_\parallel^{-2} P_\perp(k_\parallel) [1 - \cos(k_\parallel \Delta z)],$$  \hspace{1cm} (19)

with relatively simple analytical expressions for piecewise power-law spectra (see Ragot 1999, 2006c, 2018). In Equation (18), $\vec{b}_x$ and $\vec{b}_y$ are the Fourier transforms (with sliding windows) of the rescaled x- and y-components of the magnetic field $\mathbf{B}$ (i.e., $\mathbf{b} = \mathbf{B}/B_0$ where $B_0$ is the intensity of the background field) and the wavenumber $k_\alpha$ is related to the definition of the Fourier transform $\vec{b}$ with a window of finite width or, if phase correlations exist in the spectrum, to the phase-correlation scale (see Ragot 2006c).

For our purpose here of determining the correlation length, we only consider the limits of the very short and very long scales $\Delta z$. In the limit of the very short scales, such that $k_\alpha \Delta z \ll 1$, the integral expression of Equation (19) behaves as

$$\langle (\Delta r)^2 \rangle \sim 2 (\Delta z)^2 \int_{0}^{k_\alpha} dk_\parallel P_\perp(k_\parallel) = I_0(\Delta z)^2,$$  \hspace{1cm} (20)

where we have introduced

$$I_0 \equiv \frac{1}{T} \int_{0}^{\alpha_{0} + T} dt \left[ b_x^2 (t) + b_y^2 (t) \right],$$  \hspace{1cm} (21)

which is much easier to evaluate numerically than power spectra and their integral. (The spectral integrals and $I_0$ are related through the Parseval equality.) At the long scales $\Delta z > 1/k_0$, where $k_0$ is defined as the parallel wavenumber below which the spectrum becomes flat, we have (Ragot 1999)

$$\langle (\Delta r)^2 \rangle \approx 4 \int_{0}^{k_0} dk_\parallel k_\parallel^{-2} P_\perp(k_\parallel) [1 - \cos(k_\parallel \Delta z)]$$  \hspace{1cm} (22)

$$= 4 P_\perp(k_0) \int_{0}^{k_0} dk_\parallel k_\parallel^{-2} [1 - \cos(k_\parallel \Delta z)]$$  \hspace{1cm} (23)

\(^5\) The $r$ that appears here after a $\Delta$ or $\delta$ should of course not be confused with the correlation coefficient $r$ of Section 2, which does not appear in Section 3.
Long-scale $A_\perp$ of Equation (31), over $2\beta$ times the background field aligned length scale $\Delta z$, as a function of $k_0\Delta z$, $k_0$ being defined as the parallel wavenumber below which the spectrum flattens.

$$A_{s_l}(k_0, k) = 4 P_s(k_0) \left[ \Delta z \sin(k_0\Delta z) - \frac{1 - \cos(k_0\Delta z)}{k_0} \right],$$

(24)

which in the limit $k_0\Delta z \gg 1$ increases linearly with $\Delta z$,

$$(\Delta r_f)^2 \sim 2\pi P_s(k_0) \Delta z.$$  

(25)

In Equation (24), $\text{Si}$ is the Sine-integral function defined by $\text{Si}(x) \equiv \int_0^x dt \sin(t)/t$.

### 3.3. Application to Data Analysis

In terms of $I_b$ (see Equation (21)), we can write

$$k_0 P_\perp(k_0) + \int_{k > k_0} dk_0 P_\parallel(k_0) = \frac{I_b}{2},$$

(26)

and it is therefore convenient to consider the quantity

$$A_\perp \equiv \frac{\langle (\Delta r_f)^2 \rangle}{I_b \Delta z},$$

(27)

with the two simple asymptotes:

$A_\perp \sim \Delta z$ for $\Delta z \ll k_M^{-1},$  

(28)

and

$A_\perp \sim A_{\perp, \infty} = \frac{\pi \beta}{k_0}$ for $\Delta z \gg k_0^{-1},$  

(29)

where

$$\beta \equiv 1 - \frac{2}{I_b} \int_{k > k_0} dk_0 P_\parallel(k_0) \leq 1$$

(30)

is the relative magnetic power at $|k_\parallel| < k_0$.

More generally, between the asymptotes but for long $\Delta z > 1/k_0$, in the quasilinear regime of turbulence, we can rewrite Equation (24) for $A_\perp$ as

$$A_\perp \sim \frac{2\beta}{k_0} \left[ \text{Si}(k_0\Delta z) - \frac{1 - \cos(k_0\Delta z)}{k_0\Delta z} \right].$$

(31)

Figure 18 shows $A_{\perp}$ over $2\beta\Delta z$ computed from Equation (31), as a function of $k_0\Delta z$ between 0.1 and 30 in log–log scale. In the quasilinear regime, the MFL decorrelation is initiated at the parallel correlation scale $L_c = 1/k_0$ (Ragot 1999, 2006c, 2018). This is the scale where Figure 18 shows the function $A_{\perp}/\Delta z$ starting to head downward (indicated by black vertical line). The actual MFL diffusion (asymptote of Equation (29)) requires a few such correlation scales to start. It is effectively started by the scale $L_2 = 2\pi/k_0$ of the first minimum in Figure 19. The red vertical line indicates the position $k_0\Delta z = 2\pi$ of that first minimum of the slope computed in log–log space.

When analyzing the $A_\perp$ computed from SW data, because of the rapidly growing error bars at longer $\Delta z$ scales, it is more practical to determine the scale where a given value $s_{\text{lim}}$ of the log–log slope of $A_\perp$ versus $\Delta z$ is first reached. Assuming that the interval of the SW data analyzed is indeed in the quasilinear regime of turbulence, we define the length scale

$$L_\gamma(s) \equiv \frac{\gamma(s)}{k_0},$$

(32)

and show in Figure 19 the value of $s = s_{\text{lim}}$ as a function of $\gamma$. The slope $s$ becomes less than $s_{\text{lim}} = 0.1$ at approximately $L_\gamma = 5/k_0$. Still assuming quasilinearity, if for a given data interval we can determine the value of $L_\gamma$ for a given $s_{\text{lim}}$, using the relation $\gamma = \gamma(s_{\text{lim}})$ of Figure 19 we can deduce the value of the wavenumber $k_0$ on that interval, and obtain under the assumption of quasilinearity, the parallel correlation length,

$$L_{c\parallel} \equiv \frac{1}{k_0} = \frac{L_\gamma(s_{\text{lim}})}{\gamma(s_{\text{lim}})},$$

(33)

and the diffusion length,

$$L_{d\parallel} \equiv \frac{2\pi}{k_0} = \frac{2\pi L_{c\parallel}(s_{\text{lim}})}{\gamma(s_{\text{lim}})}.$$  

(34)

Before concluding that these length scales are the actual correlation and diffusion lengths, respectively, we still need to determine whether we are in the quasilinear or nonlinear regime of turbulence, and that SW cross flow is not the cause of the decorrelation. In all cases, diffusion starts on the observed scale, so the numerical value of the diffusion length is not affected, and the inferred value of the correlation length is probably close to the real correlation length (most definitely of the right order of magnitude). However, the interpretation of

6 In Figures 22, 23, 25, and 26, the uncertainties become large enough to affect the visible slope beyond the wavy vertical lines. The lines are wavy to indicate that their position is not the same for all intervals, due to variations in $V_{SW}$. 

Figure 18. Long-scale $A_{\perp}$ of Equation (31), over $2\beta$ times the background field aligned length scale $\Delta z$, as a function of $k_0\Delta z$, $k_0$ being defined as the parallel wavenumber below which the spectrum flattens.

Figure 19. Slope $s$ of the long-scale $A_{\perp}$ of Equation (31) vs. $\Delta z$, in log-log space, as a function of $k_0\Delta z = \Delta z/L_{d\parallel}$.
this length very much depends on the cause of the decorrelation.

3.4. Nonlinear versus Quasilinear: a Practical Test

A fully nonlinear statistical calculation of the mean CFD (Ragot 2006b) establishes the range of validity for the quasilinear approximation. It predicts a nonlinear diffusion of the MFLs, distinct from and slower than the quasilinear or GQL diffusion, whenever the mean CFD \( \langle (\Delta r)^2 \rangle^{1/2} \) exceeds twice the perpendicular correlation length \( L_{c_2} \) and the nonlinear length scale \( \zeta_0 \) defined below in Equation (35) is shorter than the parallel correlation length \( L_{c_1} \). When \( \zeta_0 \) is not shorter than \( L_{c_2} \), that is, at lower levels of turbulence, the quasilinear or GQL diffusion sets in first and both GQL and nonlinear diffusion coefficients are identical. In both cases, the diffusion is preceded by a supradiffusion, that is, a transport regime wherein \( \langle (\Delta r)^2 \rangle \) increases faster than \( \Delta z \). In the nonlinear regime, decorrelation and diffusion are driven by the MFL CFDs.

For nonsingular spectra,\(^7\) Ragot (2006b) showed that, whenever the nonlinear length scale \( \zeta_0 \), defined by

\[
\langle (r_{\omega} - r_0)^2 \rangle^{1/2} \equiv \sigma_{\omega} \equiv \sigma_0 = 2 \xi k_0^{-1},
\]

is less than the quasilinear parallel correlation length \( L_{c_2,QL} \approx k_0^{-1} \), a nonlinear diffusion regime sets in beyond \( \Delta z = \zeta_0 \), for which the mean square CFD, \( \sigma_{\Delta z}^2 \equiv \langle (\Delta r)^2 \rangle \), can be approximated by

\[
\sigma_{\Delta z}^2 \approx \sigma_0^2 + 2 (\Delta z - \zeta_0) D_{NL},
\]

with the nonlinear diffusion coefficient (see Figure 6 of Ragot 2006b)

\[
D_{NL} \approx 4 \int_0^{\zeta_0} d\zeta \int_0^{\min(2\sigma_{\omega}^{-1}k_{p+1})} d\kappa \parallel \times \cos(k_1\zeta) P_{\perp\text{nl},p}(k_1, \zeta).
\]

In Equation (35), the anisotropy parameter,

\[
\xi \equiv \frac{L_{c_2}}{L_{c_1}},
\]

is the ratio of the perpendicular to the parallel correlation lengths, in the quasilinear regime \( (L_{c_2} \text{ is } L_{c_2,QL}, \text{ not } \zeta_0) \).

For anisotropic elliptical turbulence, for which \( \beta^2(k) \propto C_p (k_\parallel^2 + \xi^2 k_\perp^2)^{\gamma/2} \) when \( k_{p+1}^2 < k_\parallel^2 + \xi^2 k_\perp^2 < k_{p+1}^2 \), the components of the nonlinear spectral function, \( P_{\perp\text{nl},p} \), are given by Equation (A8) of Appendix A of Ragot (2010) and the GQL mean square CFD \( \sigma_{\Delta z}^2 \) at \( \zeta < \zeta_0 \) by Equation (A13) of that same Appendix.\(^8\) Because \( \xi \) is less than \( \zeta_0 \) within the integral of Equation (37), \( \sigma_\xi^2 \) can be accurately estimated using the GQL result of Ragot (2006c or 1999). Both Equations (A8) and (A13) are closed analytical expressions. So the GQL and nonlinear mean CFDs and diffusion coefficients can be estimated from the turbulence spectra. They closely match, in Figures 5 and 6 of Ragot (2006b), the results of GQL and fully nonlinear numerical simulations of MFLs.

In terms of the parameters defined in this paper, accessible through simple data analysis, the nonlinear regime is reached if there exists a length scale \( \zeta < L_{c_2,QL} = 1/k_0 \) such that

\[
h(\zeta) > \frac{\xi}{\beta},
\]

where

\[
h(\zeta) \equiv \frac{\pi b^{1/2} (\zeta A_{\infty})^{1/2}}{2 A_{\infty}},
\]

Again, \( \beta \), the relative magnetic power below \( k_0 \), is defined in Equation (30), and the anisotropy parameter \( \xi \) in Equation (38). For each of the ACE data intervals analyzed for the year 1999, we plot \( h(\zeta) \) as function of \( \zeta \) in Figure 20 and indicate with crossed diamonds (and crosses) of matching color, the position of \( L_5 \) (and \( L_{c_2} \)). From the definitions of \( A_{\perp,\infty} \) and \( L_{c_2} \) (see Equations (29) and (32)) follows that, in the quasilinear regime,

\[
\beta = \frac{k_0 A_{\perp,\infty}}{\pi} = \frac{\gamma (s_{lim}) A_{\perp,\infty}}{\pi L_{\gamma}}
\]

can be obtained from the data analysis. The ratio \( \xi/\beta \) is shown as spiders in Figure 20 for \( \xi = 1 \) (and as thinner spiders for \( \xi = 1/3 \)). We can see from Figure 20 that unless \( \xi \) is small (\( <1/3 \)), \( h(\zeta) \) remains less than \( \xi/\beta \) up to \( L_{c_2} \) (indicated by the crosses, and dotted vertical lines) and the quasilinear condition is satisfied. It is likely that \( \xi < 1 \), on the order of 1/2, 1/3, or slightly less. One-third would be a plausible value. So we note that although we believe here that the nonlinear regime with its early decorrelation is not reached, it is very close (as would be expected from Figures 5–6 of Ragot 2006b).

Figure 21 shows for all 37 day intervals from 1999 to 2018, the product \( \beta h(\zeta_0) \) when \( \zeta_0 = L_{c_2} = L_{c_2,QL} \), that is, the value of
ξ that would put the turbulence right at the limit of the nonlinear regime. Shorter $L_c$ values tend to imply an easier condition with larger $ξ$ values, which is not really surprising since $ξ = L_c/L_c$. (The slope of the fitting red line in Figure 21 is only about $-0.22$, though.) From Figure 21 it appears likely that the nonlinear regime is reached on at least some SW intervals.

If $ξ$ were much smaller than 1 most of the time, then the nonlinear regime would be reached most of the time and the nonlinearity would be strong. So we consider here what the consequences would be for our analysis. Well within the nonlinear regime, the expressions for the measured parameters $A_{⊥\infty}$ and $L_γ$ should be rescaled by $C^{1/2}$ and $C^{-1/2}$, respectively, with $C > 1$, the turbulence level relative to that at which the nonlinear regime starts (see Ragot 2006b, below Figure 5). The new nonlinear scalings,

$$A_{⊥\infty} = \frac{\pi \beta C^{1/2}}{k_0} \quad (42)$$

and

$$L_γ = \gamma \frac{ζ_{\text{0NL}}}{k_0} = \frac{\gamma}{C^{1/2}} \quad (43)$$

imply

$$k_0 = \frac{\gamma}{L_γ C^{1/2}} \quad (44)$$

and

$$\beta = \frac{A_{⊥\infty}}{\pi ζ_{\text{0NL}} C} = \frac{γ A_{⊥\infty}}{π L_γ C} \quad (45)$$

We would therefore be deducing from the measured parameters smaller $β$ values (by a factor $C^{-1}$), making the condition (39) for nonlinearity harder to reach. So if the values of $ξ$ are indeed smaller than we believe they are (smaller than 1/3), the nonlinear regime should still be relatively “mild” or marginal, with only relatively small corrections to quasilinearity. The MFL decorrelation would be due to nonlinear CFDs within the SW rather than, e.g., strong temporal variations of the magnetic fields emerging from the lower solar corona, but on a timescale only marginally shorter.

3.5. MFL Wandering versus SW Cross Flow

Another possible interpretation of the observed diffusivity and inferred correlation length $L_c$ is that the observed MFL decorrelation is due to the SW cross flow. On the Parker spiral or background field aligned scale, $Δz = V_{SW} Δt / \cos ψ$,\(^9\) the SW flows past the spacecraft a distance $V_{SW} Δt \sin ψ = Δz \cos ψ \sin \psi$ along $x$, corresponding to

$$A_{V_{SW}} = \frac{(Δr)^2_{V_{SW}}}{l_0 Δz} = \frac{(Δz \cos ψ \sin ψ)^2}{l_0 Δz} \quad (46)$$

where $ψ$ is the angle between the Parker spiral and the radial. We compare in Figures 22 and 23 the values of $A_{⊥}$ (crosses on the $A_{⊥} = A_{⊥}(Δz)$ lines) and $A_{V_{SW}}$ (spiders on the straight $A_{V_{SW}} = A_{V_{SW}}(Δz)$ lines) at $Δz = L_c$ for the years 1999 and 2009, respectively. In the second figure, $L_c$ is often further

\(9\) In Ragot (2006e), we used $Δz = V_{SW} Δt \cos ψ$; but we now believe $Δz$ should be the local length of the Parker spiral, not the projection of $V_{SW} Δt$ on the Parker spiral direction.
away from the intersecting lines, so it is easier to read. When the spiders are well below the crosses, field-line wandering clearly dominates the SW cross flow. When the spiders are well above the crosses as happens once in the 2009 figure and twice in the 1999 figure for the highest values of $L_{c\parallel}$, SW cross flow is clearly the cause of the decorrelation and subsequent diffusion. In all other cases, both MFL wandering and SW cross flow probably contribute in some measure to the decorrelation and diffusion.

Figure 24 shows the ratios $A_{37\text{d}}/A_{c\parallel}$ for all 37 day intervals of our analysis. It is clearly $\lesssim 1$ in the vast majority of the cases, though again, for the largest values of $L_{c\parallel}$, most values of the ratio exceed one. It seems that in those few cases when MFL wandering fails to produce a decorrelation, the SW cross flow takes over and forces the decorrelation.

### 3.6. Results for the Correlation Length $L_{c\parallel}$

Figure 25 summarizes our data analysis over the entire data set, for individual $\approx 37$ day long intervals. Each panel displays the computed $A_{c\parallel}$ versus background field aligned length scale $\Delta z$ in cm, the locations of our automatically estimated $(L_{c\parallel}, A_{c\parallel})$, for $s_{\lim} = 0.1$, and a waving vertical line indicating the scale beyond which error bars become significant, at least after decorrelation has occurred. To increase the statistics in the computation of $A_{c\parallel}$, the subintervals of scale $\Delta z$ were taken at increments $\Delta z/16$.

The year averages of $A_{c\parallel}$ versus $\Delta z$ are shown in Figure 26 for the 20 years of our data analysis. Most give an $L_{c\parallel}(s_{\lim} = 0.1)$ less than about $2 \times 10^{13}$ cm, which is where the uncertainties cause $A_{c\parallel}$ to start wavering around, in the cases where decorrelation has occurred. When strong correlations subsist, the uncertainties are less. This is why we do not dismiss the few cases where $A_{c\parallel}$ continues increasing well past $\Delta z = 2 \times 10^{13}$ cm.

Figure 27 shows the distribution of correlation lengths obtained through the CFD method of this section for all successive 37 day intervals of the years 1999–2018. We note that unlike with the correlation-function methods, we obtain just one distribution of correlation lengths $L_{c\parallel}$, computed with intervals of duration $T = 37$ days. No diffusion could be observed with enough accuracy with shorter intervals (of duration $T = \delta t \times$ a power of 2, to allow for a fast computation of power spectra as well), and there is little ambiguity about the definition of the correlation length and the way to obtain it. Analysis on shorter time intervals, in most cases, would only provide lower limits for the correlation and diffusion lengths (as was the case in our earlier analysis of Ragot 2006e—see Table 1 of that paper).

The values of the correlation length $L_{c\parallel}$ range from about 0.037 to 0.3 au and, with lower probability, extend past 1 au. The distribution peaks at 0.075 au, with a secondary peak at 0.15 au. The red dotted, green dotted, and blue dotted histograms in Figure 27 show the contributions from the nonlinear regime, the quasilinear regime, and the regime dominated by the SW cross flow, respectively. In the nonlinear regime of turbulence, the computed $L_{c\parallel}$ is the $Q_{0} < 1/\xi_{0}$ of Equation (35). It tends to be shorter because the nonlinearity produced by strong MFL wandering causes an early decorrelation. The limit between nonlinear and quasilinear regimes was computed in Figure 27 with $\xi = 0.25$. For $\xi = 0.25$, we obtain similar numbers of SW intervals in the nonlinear and the quasilinear regime of turbulence. For $\xi = 0.33$, most SW intervals would be in the quasilinear regime, while for $\xi = 0.20$, most would be in the nonlinear regime. So the distributions of nonlinear versus quasilinear regimes is quite sensitive to the value of $\xi$. In the “cross-flow dominated” regime, MFL wandering fails to produce a decorrelation. Eventually, the SW cross flow at the spacecraft causes that decorrelation of the fields, but on a longer length scale, hence the tail of measured $L_{c\parallel}$ distribution.

The range of values of the parallel correlation length $L_{c\parallel}$ in Figure 27 is consistent with the $T = 37$ days correlation analysis of Section 2 (see the right panels of Figure 13 and top panel of Figure 12), the best consistency being with the $T = 37$ days analysis of the $l$-averaged data (see red histograms in upper panel of Figure 12). But we defer until Section 4 the comparison and consistency check of the correlation lengths obtained here in Section 3 through the CFD diffusivity test and earlier in Section 2 through the use of correlation functions. To conclude on the use of CFDs to evaluate a correlation length $L_{c\parallel}$, in the cases when MFL wandering dominates the cross flow, whether the observed diffusion is quasilinear or nonlinear, the MFL diffusion and correlation lengths are, in most cases with a high degree of confidence, the ones that can be deduced, with very little need for interpretation, from Figures 22, 23, or more generally, Figure 25. The measurements with automated computing of the scales $L_{c\parallel}$ and $A_{c\parallel}$ can in some cases give excessive values, but visual inspection of the figures for mean CFDs should allow one to correct and/or properly interpret these results. For instance, in the Year 2007 panel of Figure 25, the “runaway” straight green line and diamond at $10^{12}$ cm indicate a failure of MFL wandering to produce decorrelation and diffusion; this is clearly a cross-flow dominated interval. It could be worth checking for the presence of CMEs, but we find that such structures are not necessary to produce this type of CFD behavior.

In general, we find that using the CFD diffusivity test of this section to evaluate diffusion and correlation lengths in the SW is more straightforward than the methods based on correlation functions (in part because it is easier to know when convergence/diffusion has been reached). The results are easier to interpret and well suited for application to transport problems. The obvious drawback is the length of the intervals needed to observe the diffusion (five times the correlation length $L_{c\parallel}$, which is already long), and therefore, the lack of locality of the evaluated length scale.
3.7. Inferred Spectral Power at Low Frequencies

Finally, what can we learn from our data analysis about the spectral power at low frequencies? And could that information have been obtained through direct spectral analysis of the fields?
From the data analysis of the magnetic fields we obtain for each SW interval the length $L_\gamma$, where the mean CFD/Δz (or $A_\perp$) slope versus scale Δz, in log-log space, first reaches $s_{\text{lim}} = 0.1$, and the value of the asymptote $A_\perp \propto z$. We also compute the mean value $I_0$ of the cross-field magnetic energy relative to the background magnetic energy (see Equation (21)). This is also the total spectral power $2 \int_{k_{\perp 0}} P_\perp(k_\perp)$ in the $b_\perp$ and $b_\parallel$ fields (Parceval equality). The spectral power at “zero,” $2 P_\perp(k_\perp = 0)$, of these fields is the one that was needed in the early quasilinear theory approach to obtain the correlation length (Jokipii & Coleman 1968; Jokipii & Parker 1969).

From the definition of $\beta$ (the relative magnetic power at $|k_\parallel| < k_0$—see Equation (30)), we have

$$k_0 P_\parallel(k_0) = \frac{\beta I_0}{2}$$

and

$$\int_{k_{\perp 0}}^1 dk_\perp P_\perp(k_\perp) = \frac{(1 - \beta) I_0}{2}.$$  

Together with Equations (44) and (45), it leads to

$$P_\perp(k_0) = \frac{I_0 A_\perp \propto}{2\pi C^{1/2}}$$

and

$$\int_{k_{\perp 0}}^1 dk_\perp P_\perp(k_\perp) = \frac{I_0}{2} \left[ 1 - \frac{\gamma A_\perp \propto}{\pi L_\gamma C} \right].$$

where $C (>1)$, the turbulence level relative to that at which the nonlinear regime starts, is unknown but, we believe, can safely be assumed of order one (see Section 3.4). If the nonlinear regime does not start, that is, in the quasilinear regime, $C = 1$. All other quantities are known quantities of the data analysis.

To see how the data analysis can further constrain the power spectrum, let us assume that $P_\perp(k_0) = P_\perp(k_0/k_0)$ with $-1 \leq a < 0$ up to a parallel wavenumber $k_1$ and is steeper than $k_{\perp 0}^{-1}$ above $k_1$, with a spectral index $a' < -1$ (usually $a' \sim -1.5$ to $-5/3$). If $-1 < a < 0$,

$$\int_{k_{\perp 0}}^{k_1} dk_\perp P_\perp(k_\perp) \approx \frac{k_0 P_\perp(k_0)}{1 + a} \left[ \left( \frac{k_1}{k_0} \right)^{1+a} - 1 \right] + k_0 P_\perp(k_0/a') - 1$$

$$\approx \frac{\gamma I_0 A_\perp \propto}{2\pi C^{1/2} L_\gamma C^{1/2}} \left[ \frac{a' - a}{a' + 1} \left( k_1 \right)^{1+a} \right] - 1.$$

which together with Equation (50) implies

$$k_1 \approx k_0 \left[ \frac{a' + 1}{a' - a} \left[ 1 + \left( 1 + a \right) \left( \frac{\pi L_\gamma C}{\gamma A_\perp \propto} - 1 \right) \right] \right]^{1+a' \gamma}.$$  

If $a = -1$, we find instead

$$k_1 \approx k_0 \exp \left[ \frac{\pi L_\gamma C}{\gamma A_\perp \propto} - \frac{a'}{a' + 1} \right].$$

Now, let us compare these inferred $P_\perp(k_0)$ spectra to the spectra that can be computed directly by Fourier analysis of the ACE magnetic field data (see, as an example, frequency spectra of Figure 28). The spectra $P_{\text{obs}}$ computed directly by Fourier analysis of the in situ data are not function of $k_0$, the wavenumber along the background field or spiral direction, but of $k_\parallel$, the wavenumber along the radial direction. They are projected spectra, but on a direction that is at an angle $\psi \sim 45^\circ$ (on long time intervals in the ecliptic) to the one needed to estimate the mean CFDs. Furthermore, the total duration $T$ of the SW intervals needed to statistically observe the onset of CFD diffusion is so long that the fields will clearly have decorrelated, several times over, due to the spanning of magnetic fields emerged from several independent supergranulation cells back at the Sun. The effect should be a flattening of $P_{\text{obs}}$, already at frequencies of several times $1/T$ (see Figure 28). That flattening is a cross-flow effect (see also discussions in Ragot 2006d, 2006e).

In our data analysis, the diffusion onset is usually observed at or before (within a factor of a few) one tenth the total length spanned by a given SW interval, and as argued in Section 3.3, is due to decorrelation starting at a scale that is another $1/\gamma$ times shorter. So typically, the correlation times are $\tau_c \leq T/50 \leq 18$ hr. As we carefully checked in Section 3.5 (see Figures 22–24), the MFL mean CFDs at the correlation times exceed the cross flow, except in the very few cases in which the decorrelation due to MFL wandering fails to occur and cross-flow decorrelation substitutes. The decorrelation due to MFL wandering, when it occurs, is due to spectral flattening of $P_\perp(k_0)$, at frequencies $\nu_0 = \nu_{\text{SW}} \cos \psi k_0/(2\pi)$ that are too

We could not just have chosen the radial as the z-direction for our analysis because for the theory to apply, be it generalized quasilinear, nonlinear, or simply quasilinear, the z-direction has to be the direction of the background field, the magnetic field that remains in the absence of turbulence or perturbation of any kind. Due to the solar rotation, that field is along the Parker spiral, not the radial.

That is no coincidence. We have chosen the duration $T$ so as to have a statistics of at least ten successive $\Delta z$ over the interval duration at the onset of diffusion. Fewer than that and the fluctuations in the curves $A_\perp(\Delta z)$ would be too large, masking the onset. To further reduce the fluctuations in $A_\perp$, we chose an increment $\Delta z/10$ (rather than $\Delta z$) between successive $\Delta z$ segments.
low for spectral analysis on intervals of duration $\tau_c$ (or even 5 $\tau_c$) to reach. Using the entire interval of duration $T$ to compute a spectrum $P_{\text{obs}}$ and compare the result with the spectra that can be deduced from the CFD analysis is misleading because a proper comparative study of the spectra should compute $P_{\text{obs}}$ from intervals of durations $T/10$, or even $T/50$, and average the results. But then the flattening frequency would not be reached, and there would not be much to compare.

To conclude on this point, $k_0$ and $P_{\perp}(k_0)$ cannot be computed by spectral analysis of the data. Any spectral analysis of SW data intervals with sufficient duration to reach the relevant low frequencies is doomed by transverse decorrelation, due to the rotation of the Sun. This explains why the correlation length $\sim10^{-2}$ au obtained early on from the direct computation of the power at “zero” frequency (Jokipii & Coleman 1968; Jokipii & Parker 1969) is a serious underestimate of the parallel correlation length $L_{c||}$.

### 4. Reconciling the Results

We have so far more or less considered the two approaches for estimating the correlation lengths separately, focusing in Section 2 on the correlation functions approach and in Section 3 on the CFD diffusivity test. The “third” method involving the spectral power at “zero” frequency (Jokipii & Parker 1968, 1969) is addressed in Section 3 for the obvious reason that the CFD diffusivity test is based on a theory that generalizes the original work by Jokipii and Parker, and both theories relate the mean CFDs to the projected power spectra (even if in this paper, we have worked our way around the power spectra to estimate the correlation length in a more reliable way, precisely due to the difficulty of evaluating the power at “zero” frequency). The spectral power at “zero” frequency method and the CFD diffusivity-test method both deal with CFDs and diffusivity of MFLs. They estimate the correlation length of the MFLs, which relate to the magnetic field integrated over a path. By contrast, the correlation functions approach is generally applied to the magnetic field itself, not its integral along a path.

We could argue here that energetic charged particles propagating in the SW feel the cumulative effect of the SW magnetic fields along their path, that their path to first order follows the MFLs, and that therefore the MFL correlation length is most relevant to the transport of energetic charged particles, whereas the usefulness of the correlation length of the “instantaneous” field in helping understand/describe the transport of these particles is somewhat more removed. But here our focus is on reconciling the results of both methods, the method based on the correlation functions and the method based on the CFD diffusivity test (the issue with the spectral power at “zero” frequency having been solved in Section 3.7). So we just do that below, reconcile the results of both methods, with the understanding that the results of the two methods generally are the correlation lengths or times of two different quantities.

In Section 2.3, we expressed in Equation (5) the correlation function $C_A(\tau)$ of the field $A$ as a ratio of integrals, over frequency, of the power spectrum $P_A(\nu)$. From this expression of Equation (5), we can estimate the correlation time $\tau_{c,A}$ of a field $A$ in the limit of an infinite duration $T$ as follows:

$$\tau_{c,A} = \int_0^{+\infty} d\tau C_A(\tau) = \int_0^{+\infty} d\tau \frac{\int d\nu e^{2\pi i \nu \tau} P_A(\nu)}{\int d\nu P_A(\nu)}$$

$$= \frac{\int d\nu P_A(\nu) \int_0^{+\infty} d\tau e^{2\pi i \nu \tau}}{\int d\nu P_A(\nu)}$$

$$= \frac{\int d\nu P_A(\nu) \delta(\nu)}{2\int d\nu P_A(\nu)} = \frac{1}{4} \frac{P_A(0)}{\int_{\nu>0} d\nu P_A(\nu)}. \quad (55)$$

The correlation-function approach is applied in Section 2 to the $B_\| \times$ component of magnetic field measured in the SW on board ACE. So, if the correlation-function method is applied properly, in the limit of the very long interval durations $T$, it should produce correlation times

$$\tau_{c,B_\|} = \frac{1}{4} \frac{P_{B_\|}(0)}{\int_{\nu>0} d\nu P_{B_\|}(\nu)}. \quad (56)$$

On the other hand, the field-line correlation length along the Parker spiral of Section 3 (see Equations (33) and (43) for the quasilinear and nonlinear regimes)

$$L_{c||} = \frac{L_\gamma}{\gamma} = \frac{1}{k_0 C_{\gamma}^{1/2}} = \frac{V_{\text{SW}} \cos \psi}{2\pi \nu_0 C_{\gamma}^{1/2}}, \quad (57)$$

corresponds to a field-line correlation time

$$\tau_{c,MFL} = \frac{L_{c||} \cos \psi}{V_{\text{SW}}} = \frac{\cos^2 \psi}{2\pi \nu_0 C_{\gamma}^{1/2}}. \quad (58)$$

in the nearly radially flowing SW. The field lines are wandering across the background magnetic field $B_0$ in a two-dimensional space, the ($x$, $y$) plane. In order to check the consistency of the results of both methods, we need to reduce the wandering of the field lines to just one dimension, along the $x$-line. This can be very easily done by substituting the spectrum $P_{B_\|}$ and field’s Fourier transform $\hat{B}_0$ for $P_{\perp}$ and $\hat{B}_0$, and repeating the exact same computations to obtain $A_{\perp}$, $A_{\infty}$, and $L$, in place of $A_{\perp}$, $A_{\infty}$, and $L$, (see Figure 25). We are now using the same fields and projected power spectra in both methods, within the factor $\hat{B}_0$, which cancels due to the ratio in Equation (56).

Since $\nu_0$ is the frequency below which the projected spectrum is flat, $P_{B_\|}(0) = P_{B_\|}(\nu_0)$ and

$$\frac{\tau_{c,B_\|}}{\tau_{c,MFL}} = \frac{C_{\gamma}^{1/2} \pi \beta}{2 \cos^2 \psi}. \quad (59)$$

where $\beta = \gamma A_{\infty}/(\pi L C)$ (see Equation (45), with $\gamma = 5$ for a start of the diffusion estimated at a log–log slope $\alpha_{\text{lim}} = 0.1$) is the relative magnetic power at $|\nu| < \nu_0$ or $|k_\parallel| < k_0$ (see also Equation (30)), and $C (\geq 1)$ is the turbulence level relative to that at which the nonlinear regime starts. $C$ is unknown but again, we believe, can safely be assumed to be on the order of one (see Section 3.4). Expressed as a function of “known”
Figure 29. PDF of the ratio $\frac{\pi \beta}{2 \cos^2 \psi}$, where $\beta$ is the relative magnetic power at $\nu < \nu_0$ (where the spectrum becomes flat). When multiplied by $C_{1/2}$, $\frac{\pi \beta}{2 \cos^2 \psi}$ gives the theoretical ratio $\tau_{c, B_0}/\tau_{c, MPL}$ in the limit of the very long times $T$ (see Equation (59)).

Figure 30. Ratios $\tau_{c, B_0}/\tau_{c, MPL}$ over $\frac{\pi \beta}{2 \cos^2 \psi}$ as functions of the field-line correlation times $\tau_{c, MPL}$ in hours, for $T = 37$ hr. The histograms of the ratios and of $\tau_{c, MPL}$ are shown on the right and top of the main panel. The ratios’ median is at 1.23 (red solid line) and the 10th (20th) and 90th (80th) percentiles at 0.81 (0.93) and 1.70 (1.48) are indicated by the red dotted (dashed) lines, while the black solid and dashed lines show 1, and 0.5 and 2, respectively. This figure demonstrates the consistency of the correlation times estimated through the correlation-function method and through the CFD diffusivity-test method.

Figure 31. Same as Figure 30, but for the ratios $\tau_{1/e, B_0}/\tau_{c, MPL}$ over $\frac{\pi \beta}{2 \cos^2 \psi}$. The median and 10th (20th) and 90th (80th) percentiles are at 0.81, 0.39 (0.49), and 1.29 (1.12), respectively.

The values of $\frac{\pi \beta}{2 \cos^2 \psi}$ range from approximately 0.05 to 1. Their distribution is shown in Figure 29.

Figure 30 shows the ratio $\tau_{c, B_0}/\tau_{c, MPL}$ over $\frac{\pi \beta}{2 \cos^2 \psi}$ as a function of the field-line correlation time $\tau_{c, MPL}$ in hours, with the histograms of each quantity on the top and right of the main panel. Here in all our comparisons, $T = 37$ days, to ensure convergence (see Sections 2.2.1 and 2.3). The points with $\tau_{c, MPL}$ greater than 80 hr (for which field-line wandering failed to produce a diffusion) or $\tau_{c, B_0}/\tau_{c, MPL} < 0.1 \frac{\pi \beta}{2 \cos^2 \psi}$ (3 points out of 180) are excluded from the right-panel histogram computation and other statistics. The median at (1.23) is indicated by the red solid line, and the 10th (20th) and 90th (80th) percentiles at 0.81 (0.93) and 1.70 (1.48) by the red dotted (dashed) lines, while the black solid and dashed lines show 1, and 0.5 and 2, respectively. The histogram peaks between the median and one. The ratio $\left(\frac{\tau_{c, B_0}}{\tau_{c, MPL}}\right)\left[\frac{\pi \beta}{2 \cos^2 \psi}\right]$ remains in a narrow band around a median of 1.23, on the order of one but larger than one, which may be due to (weak) nonlinearities (see factor $C_{1/2}$ in Equation (59)). Variations in $C_{1/2}$ may also contribute to the spread. Figure 30 clearly demonstrates the consistency of the correlation times estimated through the correlation-function method and through the CFD diffusivity-test method. The match is surprisingly good given the large uncertainties of the correlations and mean CFDs at the long scales. There is a slight tendency for the ratio to decrease with increasing $\tau_{c, MPL}$. This may be due to the increasing contribution of the tails or longer scales in the correlation functions, which can easily be underestimated due to the large uncertainties at the longer scales. The nonlinearity factor $C_{1/2}$ should also be higher at lower $\tau_{c, MPL}$ (see Figures 21 and 27).

Figure 31 shows the ratio $\tau_{1/e, B_0}/\tau_{c, MPL}$ over $\frac{\pi \beta}{2 \cos^2 \psi}$ as a function of the field-line correlation time $\tau_{c, MPL}$. Here, the spread is greater and the decreasing trend with increasing $\tau_{c, MPL}$ much more pronounced. That trend is most likely due to the increasing contribution of the tails or longer scales in the correlation functions, which $\tau_{1/e, B_0}$ does not take into account.

Figure 32 shows the ratio $\tau_{c, scale-av, B_0}/\tau_{c, MPL}$ as a function of the field-line correlation time $\tau_{c, MPL}$. The spread is even greater and the decreasing trend still present. Although the scale-averaged $B_0$ correlation time $\tau_{c, scale-av, B_0}$ obtained through the correlation-function method approximates the field-line correlation time (because it is computed from the field averaged over the scale $\tau$, the much lower statistics resulting from the scale averaging makes the determination of $\tau_{c, scale-av, B_0}$ much more imprecise, with little reliable contribution from the larger scales. In this case where the statistics are so low, it may help to estimate the correlation time through the exponential approximation. Figure 33 shows the ratio $\tau_{1/e, scale-av, B_0}/\tau_{c, MPL}$ as a function of $\tau_{c, MPL}$. The points are much closer together at short $\tau_{c, MPL}$ (when there is little tail contribution past $\lambda_{1/e}$ and the $\lambda_{1/e}$ approximation is relatively good), but at higher $\tau_{c, MPL}$, the decrease is even stronger, as should be expected due to the neglect of the correlation tails in the $\lambda_{1/e}$ approximation. The stronger slope of $\tau_{1/e, scale-av, B_0}/\tau_{c, MPL}$ supports the existence of a tail in the $\tau_{c, scale-av, B_0}$ distribution, as observed for the red solid line distribution in the top panel of Figure 12.
The ratios $\tau_{\alpha,B_x}/\tau_{\text{MFL}}$ over $\pi\beta/(2 \cos^2 \psi)$ for significance levels $\alpha = 0.001$, 0.01, and 0.05 (not shown) also have a low spread at a given $\tau_{\text{MFL}}$, comparable to $\tau_{B_x}/\tau_{\text{MFL}}$ over $\pi\beta/(2 \cos^2 \psi)$, but with a downward trend at higher $\tau_{\text{MFL}}$, probably due again to the fact that the estimates of $\tau_{\alpha,B_x}$ at small $\alpha$ do not account for the correlation tails. The medians and percentiles sharply increase past $\alpha = 0.1$, but due to the large uncertainties in the correlation estimates at longer scales, we are not sure how accurate our estimates of medians and percentiles are for larger $\alpha$. We note that for $\alpha = 0.01$, the ratio tends to be of order 2, corresponding to about 1.5 for the ratio $\tau_{\alpha,B_x}/\tau_{B_x}$, which is also evident from comparing Figures 12 and 13. If correct, this would mean that the fields are still correlated at a 0.01 significance level on one-and-a-half times the correlation timescale $\tau_{B_x}$. A test run with a simple synthetic field of known correlation time confirms this result, so even though we found the result rather surprising at first, this may be true. On the reassuring side, we find from year-average estimates that the correlations drop to a 0.25 significance level by about 4 $\tau_{B_x}$, and about a 0.4 significance level by about 6 $\tau_{B_x}$, which may already be sufficient to observe diffusion (see also footnote 4 in Section 2.2.4). So, the correlation timescale is a timescale over which correlations are very strong, but beyond a few of which the correlations sharply drop.

Figure 34 shows the histograms of the 1D MFL correlation time $\tau_{\text{MFL}}$ (black solid line) and of $\tau_{\text{MFL}}$ times $\pi\beta/(2 \cos^2 \psi)$ (black dotted line), that is, of the field’s correlation time $\tau_{B_x}$ predicted from the CFD diffusivity-test method for $C = 1$, approximately matching the histograms of $\tau_{B_x}$ over $\pi\beta/(2 \cos^2 \psi)$; (the 1D MFL correlation time $\tau_{\text{MFL}}$ predicted through the correlation-function method for $C = 1$; red solid line) and of $\tau_{B_x}$ (red dotted line), respectively. Figure 35 shows a similar match for the histograms of the correlation lengths, instead of the correlation times. Figures 34 and 35 demonstrate consistency between the results of the two methods, provided that the correlations are computed on data intervals long enough for convergence (here, $T \approx 37$ days).

As a reminder of the critical effect of the interval duration $T$ (see Figures 3–7, 14–17 and Sections 2.2.1 and 2.3), Figure 36 presents here the $\tau_{C,B_x}$ histograms computed for our entire data set and a series of interval durations $T$ between 1.75 hr and
order of magnitude. At \( T = 1.75 \) hr, the correlation times are underestimated by over an order of magnitude. So yes, the correlation-function method has converged. That said, the correlation-function method with \( T = 28 \) hr, is consistent with Figures 5–7, 16, and 17.

This \( V_{SW} \)-dependence may explain some of the yearly variations in the \( (T = 224) \) hr correlation times, with a clear drop in the year 2003 (by a factor of \( \approx 1.7 \) from the year 2002, to a value of \( \approx 1.9 \) hr), when months of high-speed SW streams severely affected the yearly distribution of wind speeds (compare the blue dotted line with the red solid line in Figure 38). In addition to the \( V_{SW} \) effect, the variations in the correlation times may also correlate with the solar cycle (see Figure 39). It is apparent for \( T = 224 \) hr (red dots) but not \( T = 28 \) hr (orange dots). Due to lower statistics (only nine intervals per year), the solar-cycle effect is also much less clear for \( T = 37 \) days (not shown), even though the \( V_{SW} \) effect is still strong. The variations with solar cycle (by a factor of \( \approx 2 \); the 10th, 20th, 80th, and 90th percentiles fluctuate by a similar factor) observed for long \( T \) may contribute to the width of the distributions in Figures 12 and 13 (though not in the lower panel of Figure 12). The observation of a solar-cycle effect in the field-component correlation time computed with \( T = 224 \) hr \( \approx 9.3 \) days, but not (or hardly) for \( T = 28 \) hr, is consistent with previous results by Wicks et al. (2010) and Engelbrecht & Wolmarans (2020).

5. Conclusion

The estimate of the SW magnetic fields’ parallel correlation length, be it from the correlation functions of the measured...
fields or from the fields’ spectral power at “zero” frequency, have long pointed toward short values on the order of 0.01 au. However, evaluation of the MFL spreading or mean CFDs failed to show the decorrelation and resulting diffusion at the “expected” length scales, pointing instead toward values for the parallel correlation length along the Parker spiral on the order of 0.1 au or more. In an effort to understand this order-of-magnitude discrepancy and reconcile the approaches using correlation functions and diffusivity test of the mean CFDs, we have applied here both approaches, with renewed attention to the “details” as well as the broader sense of the calculations, to a large (20 years long) set of magnetic field and flow data from the ACE spacecraft.

Our first finding, and possibly the most impactful in solving/reducing the discrepancy, is that the correlation functions can be severely affected by the length $L$ of the SW intervals chosen to compute them, if $L$ does not exceed the correlation length $\lambda$ by a factor large enough (see Figures 3–7, 14–17, 36, and 37). It appears that one day, the longest interval duration $T = L/V_{\text{SW}}$ of the correlation studies cited in the Introduction, causes the correlations to drop prematurely. It is too short to properly compute correlation lengths on the order of 0.1 au in the SW (Section 2.2.1). For $T \sim 2$ hr, the computed correlation lengths can be underestimated by over an order of magnitude (Figures 4, 6, 7, 14, 15, 17, 36, and 37). These findings are confirmed by simulations and analytical estimates of the correlations (Section 2.3; Figures 14–17).

The correlation length of a field $B_\lambda$ is classically defined by the integral, $\lambda_{c} \equiv \int dl \ C_B(l)$, of the correlation function $C_B(l)$ over the lengths $l$. It is customary among physicists to expect correlations to decrease exponentially with the scale $l$ and to seek the characteristic length scale $\lambda$ of that decrease, either by fitting an exponential to the correlation function, or by searching for the scale $\lambda_{c}$ where the correlations reach $1/e$, instead of integrating the correlations over $l$ to obtain $\lambda_{c}$. If the correlations fit $e^{-l/\lambda_{c}}$, then $\lambda_{c} = \lambda = \lambda_{c}$. In the SW, however, the correlations do not drop exponentially (see Figures 5–7, 10, 11). So $\lambda_{c}$ could be a poor substitute for $\lambda_{c}$. Because much of the nonexponential behavior of the correlation functions occurs before they reach $1/e$, however, the length $\lambda_{c}$ appears to be a fairly good substitute for $\lambda_{c}$, most of the time (Figure 12). In very low statistics, we suggest that it may sometimes even produce better results than the integral $\lambda_{c}$, which could be affected by the large uncertainty of the values of $C_B(l)$ at larger $l$. However, large-scale nonexponential correlations could also affect the values of $\lambda_{c}$. This is supported by the greater slopes in Figures 31 and 33 relative to 30 and 32, respectively. For high enough resolution data (no scale averaging), the issues for the integral $\lambda_{c}$ are resolved by increasing $T$.

Beyond the choice of interval duration and the difficulty in obtaining sufficient statistics for the correlations at high $l$ to accurately estimate (or simply capture the main contribution to) the integral $\lambda_{c}$, we also noted (and explored for a backup method or consistency check) the statistician’s different approach to the use of correlations. Recalling in Section 2.1 the statistician’s basic principle that a correlation coefficient $r$ cannot be used directly to indicate the degree of correlation, but that the interpretation of $r$ critically depends on the number $N$ of data points or observations used to compute $r$, we proceed in Section 2.2.4 to compute the length scales $L_{\alpha}$ at which the fields are known to be correlated at a significance level $\alpha$. The length scales $L_{\alpha}$ are given in Figures 10 and 11 by the scales where the correlation functions $C_B(l)$ intersect the critical values $r_{x}(\alpha, N = L/l)$ (see Figures 1 and 2) of the correlation coefficient. Their distributions are presented in Figure 13 for interval durations of 9 and 37 days, with and without preaveraging the magnetic fields on the scale $l$. The results for $L_{\alpha}$ emphasize how strong the correlations are on the correlation length or timescale, and even up to one-and-a-half times or twice that scale. They emphasize that the correlation timescale is the timescale over which correlations are very strong, but beyond a few of which the correlations sharply drop. They otherwise mostly confirm/support our other results, with mostly the same drawback as $\lambda_{c}$ when limited to one small value of $\alpha$.

In Section 3, we turned our attention to the other approach, which consists of computing the mean CFDs $\langle [\Delta x]^{2} + [\Delta y]^{2} \rangle^{1/2}$ of the SW MFLs on intervals of duration $T$ as functions of the background field aligned length scale $\Delta z$, and testing for the diffusivity of these CFDs to find, first, the diffusion scale, and then the correlation scale, $L_{\varsigma}$, which is a fraction of the diffusion scale. Taking advantage of the large ACE data set, we are able to choose $T$ long enough (~37 days) to actually observe the diffusion start and compute large statistics of diffusion and correlation lengths. Using the generalized quasilinear theory in a simplified form that focuses on the very short and long scales (Section 3.2), we identified in Section 3.3 the rescaled quantity $A_{\perp} \equiv ([\Delta x]^{2} + [\Delta y]^{2})/\langle \Delta z \rangle$ as optimum to obtain the diffusion/correlation length from the in situ fields’ measurements, with the simple asymptotes $\Delta z$ and $\pi^{2}L_{\varsigma}$ of Equations (28) and (29) at the very short and the very long scales, where $\beta$ is the relative magnetic power at $|k_{i}| < k_{0}$, the parallel wavenumber below which the power spectrum is flat, and $L_{b}$ is the total magnetic energy relative to background. By using the relation between the logarithmic slope $s$ of $A_{\perp}$ versus $\Delta z$ and the ratio $\gamma$ of the length scale $L_{\varsigma}$ to $L_{\varsigma}$ in the quasilinear regime (see Figure 19), we are able to make accurate determinations of $L_{\varsigma}$ by first determining $L_{\gamma}$ for $s = 0.1$ and then $L_{\varsigma} = L_{\gamma}/\gamma$ with $\gamma = \gamma(s = 0.1) \approx 5$.

We further translated the test for the nonlinearity of the turbulence regime (Ragot 2006b) in terms of the new quantity $A_{\perp}$ (see Equations (39)–(40)), and applied the test to the ACE data set (see Figures 20 and 21, Section 3.4). We find that if the ratio $\xi \equiv L_{\perp}/L_{\varsigma}$ of the perpendicular to the parallel correlation length is small enough (typically $<1/3$), the regime is nonlinear rather than quasilinear. This changes the interpretation of the measured lengths. The diffusion is nonlinear rather than quasilinear (with a lower diffusion coefficient), and the measured correlation length is $L_{\varsigma} < 1/k_{0}$ rather than $L_{\varsigma} = 1/k_{0}$. However, by introducing the nonlinear scalings for the nonlinear correlation scale $L_{\varsigma}$ and diffusion coefficient $A_{\perp}$, we find that if the nonlinear regime is reached, it should still be relatively marginal, with only small corrections to quasilinearity. The MFL decorrelation would be due to the nonlinear CFDs within the SW rather than, e.g., strong temporal variations of the magnetic fields emerging from the lower solar corona, but on a timescale only marginally shorter.

We also compared the effects of MFL wandering and SW cross flow in Section 3.5. We find that MFL wandering usually dominates or, at the very least, contributes substantially to the mean CFDs (see Figures 22–24). There are a few cases, however, where MFL wandering fails to produce decorrelation and SW cross flow forces the decorrelation at much longer $L_{\varsigma}$, accounting for the high tail of the $L_{\varsigma}$ distribution in Figure 27.
Our complete data analysis for the correlation length \( L_{c,\parallel} \), over 20 yr of the ACE data set, shows values for \( L_{c} \) ranging from about 0.037 to 0.3 au and, with much lower probability, extending past 1 au (see Figure 27, Section 3.6). The distribution of \( L_{c} \) peaks at 0.075 au, with a secondary peak at 0.15 au, so overall, we find parallel correlation lengths consistent with but somewhat shorter than our earlier study of a couple of SW streams suggested. The first and second peaks are likely due to the nonlinear and quasilinear regimes of MFL wandering, respectively.

Having considered the two approaches for estimating the correlation lengths or times separately, focusing in Section 2 on the correlation-function method and in Section 3 on the CFD diffusivity-test method (and the closely related power-at-zero-frequency method), we noted in Section 4 that the results of the two methods generally are the correlation lengths and times of two different quantities. While the diffusivity-test method estimates the correlation length of the MFLs, which relate to the magnetic field integrated over a path, the correlation-function method is generally applied to the magnetic field itself, not its integral along a path. Rescaling the \( \tau_{B,\parallel} \) and other \( B_{\parallel} \) correlation-time estimates made with the correlation-function method by the ratio \( C_{\parallel}^{1/2}\pi\beta/(2\cos^{2}\psi) \) (Figure 29) of the (long-\( T \)) theoretical correlation time of \( B_{\parallel} \) over that of the MFLs (\( \tau_{\text{MFL}} \)), we demonstrate that the correlation times obtained by the two methods are very much consistent (Figures 30–35; already without accounting for nonlinearities, i.e., for a constant \( C = 1 \)), provided that the correlation-function method is applied on data intervals long enough (Figures 3–7, 14–17, 36, 37). When applied to scale-averaged \( B_{\parallel} \) fields, the correlation-function method produces, without a need for rescaling, correlation times consistent with those of the CFD diffusivity-test method, but which are highly imprecise due to the reduced statistics resulting from the scale averaging (Figures 32, 33).

The scales are consistent, but the methods do not quite equivalent. If the goal is to estimate the correlation length or time of MFLs, the CFD diffusivity-test method is much preferred because it is the one that also gives easy access to estimates of \( \beta \) (Equation (45)), needed for the conversion from \( \tau_{B,\parallel} \) to \( \tau_{\text{MFL}} \). If more local values of the correlation length or time are needed, then the correlation-function method is needed because it may be applied, with relatively good results, to data intervals that are a few times shorter than the CFD diffusivity-test method. Doing so allows one to study the correlation-time dependence on the SW speed and solar cycle (Figures 38 and 39).

Remembering that the purpose of knowing the correlation length is often to apply its value to the modeling of transport by diffusion, it would seem appropriate enough to make sure that diffusion does indeed occur on the scale of a few of the estimated correlation lengths. One sure advantage, in that regard, of defining the correlation length through the diffusivity test of Section 3 is that by the very definition of \( L_{c,\parallel} \), we know that diffusion starts at a few \( L_{c} \) (\( \approx 5 \) for a 0.1 logarithmic slope, \( 2\pi \) for a stricter zero-slope diffusion). Whereas the use of \( L_{c,\parallel} \) as defined in Section 3, seems most appropriate in problems of transport, the scales of Section 2, in particular \( L_{o,\parallel} \), would appear more useful when dealing with questions of fields’ predictability or the predictive power of SW upstream measurements. Ultimately, \( L_{c} \) and \( L_{o} \) may be answers to two distinct questions, tools to be used in two distinct problems. One may also argue that energetic charged particles propagating in the SW feel the cumulative effect of the SW magnetic fields along their path, that their path to first order follows the MFLs, and that therefore the MFL correlation length \( L_{c} \) is most relevant to the transport of energetic charged particles, whereas the usefulness of the correlation length \( \lambda_{c,\parallel} \) of the “instantaneous” field in helping understand/describe the transport of these particles is somewhat more removed.

Finally, the Parker spiral projected power spectra at low frequency, needed by the theories for the calculations of the mean CDFs and parallel correlation length \( L_{c,\parallel} \), can be inferred from our CDF data analysis (Section 3.7). They flatten at a lower frequency and have a higher power at “zero” than the spectra obtained through the direct spectral analysis of the fields (see Figure 28). The wavenumber \( k_{0} \) and “zero frequency” power \( P_{\perp}(k_{0}) \) cannot be obtained by spectral analysis of the data obtained at 1 au at a spacecraft that is practically stationary relative to the Sun. Any such spectral analysis on SW data intervals of sufficient duration to reach the relevant low frequencies is doomed by the solar rotation, which causes a spectral flattening by transverse decorrelation. This explains why the correlation length \( \sim 10^{-2} \) au obtained early on from the direct computation of the power at “zero” frequency (Jokipii & Coleman 1968; Jokipii & Parker 1969) is a serious underestimate of the MFL parallel correlation length \( L_{c,\parallel} \).

This research was funded by NASA under the grant 80NSSC20K0696 of the Heliophysics Guest Investigators-Open Program. ACE MFI and SWEPAM data were downloaded from the Caltech website at http://www.srl.caltech.edu/ACE/ and the CDA website at http://cdaweb.gsfc.nasa.gov. We thank NASA, SPDF, and the PIs Dr. N. F. Ness and Dr. D. J. McComas for making these data sets available to the broader community. We also thank a thorough anonymous reviewer for motivating us to expand our argument about the \( T \)-dependence of the correlations.

**Appendix**

**Relation between Correlation Function and Power Spectrum**

Here, we relate the correlation function \( C_{\perp}(\tau) \) at the timescale \( \tau \) on an interval of duration \( T \), defined by

\[
C_{\perp}(\tau) \equiv \frac{\int_{0}^{T-\tau} dt \left[ A(t) - \overline{A} \right] \left[ A(t+\tau) - \overline{A} \right]}{\int_{0}^{T} dt \left[ A(t) - \overline{A} \right]^{2}},
\]

(A1)

to the Fourier power spectrum \( P_{\perp}(\nu) \).

With the Fourier and inverse Fourier transforms of a function \( f \) defined for a sliding window \( \sigma \) of width \( T \) by

\[
\mathcal{F}[f(t)]_{\nu} = \tilde{f}(\omega) \equiv \frac{1}{2\pi} \int_{0}^{T} dt \, f(t) \, \sigma(t) \, e^{-i\omega t} \quad \text{(A2)}
\]

\[
f(t) \sigma(t) \equiv \int d\omega \tilde{f}(\omega) \, e^{i\omega t}, \quad \text{(A3)}
\]

and using the convolution or Faltung theorem, which states that the Fourier transform of the convolution product of two functions is the product of the Fourier transforms of these two functions (with the convolution product \( \times \) defined with the same factor of \( 1/(2\pi) \) as the Fourier transform), we can write...
for the function $A - \overline{A}$,

$$P_A(\nu) = 2\pi P_A[\nu(\omega)] = \frac{(2\pi)^2}{T} |\mathcal{F}[A(t) - \overline{A}]|_2^2$$

$$= \frac{(2\pi)^2}{T} \mathcal{F}[\{A(t) - \overline{A}\} \times \{A(t) - \overline{A}\}]$$

$$= \frac{1}{T} \int d\tau \sigma(\tau) e^{-i2\pi\nu\tau} \int_0^{T-\tau} dt \{A(t) - \overline{A}\}[A(t + \tau) - \overline{A}] . \quad (A4)$$

From the Parceval equality, we also have

$$\int d\nu P_A(\nu) = \frac{1}{T} \int dt \sigma^2(t)|\{A(t) - \overline{A}\}|^2 . \quad (A5)$$

Taking the ratio of the above equations leads to

$$\frac{P_A(\nu)}{\int d\nu P_A(\nu)} = \int d\tau \sigma(\tau) e^{-i2\pi\nu\tau} C_A(\tau) = 2\pi \tilde{C}_A(\omega), \quad (A6)$$

and its inverse Fourier transform,

$$\sigma(\tau) C_A(\tau) = \int d\omega e^{i\omega\tau} \tilde{C}_A(\omega)$$

$$= \int d\omega e^{i\omega\tau} \frac{1}{2\pi} \int d\nu P_A[\nu(\omega)]$$

$$= \int d\nu e^{i2\pi\nu\tau} P_A(\nu) \int d\nu P_A(\nu). \quad (A7)$$

References

Balasubramaniam, K. S., & Henry, T. W. 2016, SoPh, 291, 3123
Bevington, P. R. 1969, Data Reduction and Error Analysis for the Physical Sciences (New York: McGraw-Hill)
Clette, F., Cliver, E. W., Lefèvre, L., et al. 2016, SoPh, 291, 2479
Crooker, N. U., Siscoe, G. L., Russell, C. T., & Smith, E. J. 1982, JGR, 87, 2224
Dasso, S., Milano, L. J., Matthaeus, W. H., & Smith, C. W. 2005, ApJ, 635, L181
Engelbrecht, N. E., & Wolmarans, C. P. 2020, AdSpR, 66, 2722
Jokipii, J. R., & Coleman, P. J., Jr. 1968, JGR, 73, 5495
Jokipii, J. R., & Parker, E. N. 1968, PhRvL, 21, 44
Jokipii, J. R., & Parker, E. N. 1969, ApJ, 155, 777
McComas, D. J., Bame, S. J., Beker, P., et al. 1998, SSR, 86, 563
Matsui, H., Farrugia, C. J., & Torbert, R. B. 2002, JGRA, 107, 1355
Matthaeus, W. H., & Goldstein, M. L. 1986, PhRvL, 57, 495
Matthaeus, W. H., Goldstein, M. L., & King, J. H. 1986, JGR, 91, 59
Matthaeus, W. H., Gray, P. C., Pontius, D. H., Jr., & Bieber, J. W. 1995, PhRvL, 75, 2136
Matthaeus, W. H., Dasso, S., Weygand, J. M., et al. 2005, PhRvL, 95, 231101
Matthaeus, W. H., Dasso, S., Weygand, J. M., Kivelson, M. G., & Osman, K. T. 2010, ApJL, 721, L10
Pugh, E. M., & Winslow, G. H. 1966, The Analysis of Physical Measurements (Reading, MA: Addison-Wesley)
Ragot, B. R. 1999, ApJ, 525, 524
Ragot, B. R. 2001, ICRC (Hamburg), 27, 3293
Ragot, B. R. 2006a, ApJ, 642, 1163
Ragot, B. R. 2006b, ApJ, 644, 622
Ragot, B. R. 2006c, ApJ, 645, 1169
Ragot, B. R. 2006d, ApJ, 647, 630
Ragot, B. R. 2006e, ApJ, 651, 1208
Ragot, B. R. 2010, ApJ, 715, 959
Ragot, B. R. 2018, ApJ, 868, 35
Richardson, J. D., Dashevskiy, F., & Paularena, K. I. 1998, JGR, 103, 14,619
Richardson, J. D., & Paularena, K. I. 2001, JGR, 106, 239
Ruffolo, D., Chuychai, P., & Paularena, W. H. 2006, ApJ, 644, 971
Smith, C. W., L’Heureux, J., Ness, N. F., et al. 1998, SSR, 86, 613
Weygand, J. M., Matthaeus, W. H., Dasso, S., et al. 2009, JGRA, 114, A07213
Weygand, J. M., Matthaeus, W. H., Dasso, S., & Kivelson, M. G. 2011, JGRA, 116, A08102
Wicks, R. T., Owens, M. J., & Horbury, T. S. 2010, SoPh, 262, 191
Zhao, L.-L., Adhikari, L., Zank, G. P., Hu, Q., & Feng, X. S. 2018, ApJ, 856, 94