Solar energetic electron event characteristics as inferred from spacecraft and ground-based observations

Master's thesis
University of Turku
Physics
2024
BSc Aleksi Yli-Laurila
Examiners:
Prof. R.V.
Doc. N.D.

The originality of this thesis has been checked in accordance with the University of Turku quality assurance system using Turnitin Originality Check service.

UNIVERSITY OF TURKU
Department of Physics and Astronomy

Yli-Laurila, Aleksi Solar energetic electron event characteristics as inferred from spacecraft and ground-based observations

Master's thesis, 51 pages, 3 appendices
Physics
November 2024

Massive eruptions of energy and particles from the surface of the Sun are obvious signs that the star is active. Forecasting these eruptions, however, is very difficult but of interest because of the damage that high energy particles can do to spacecraft, astronauts as well as equipment on the surface of Earth. The Sun is continuously monitored with, e.g., white-light imaging instruments such as LASCO on-board SOHO, and ground-based radio frequency instruments like the telescopes in the RTSN network. This thesis is a statistical study of solar energetic electron events, observed mainly by Solar Orbiter, combined with white-light observations of coronal mass ejections (CMEs) and ground-based observations of solar radio bursts. Solar radio bursts are classified into different types, such as type II and III bursts. They are bursts of electromagnetic radiation in the radio domain produced by electrons travelling in the solar atmosphere. The main science question is differentiating the acceleration mechanisms of solar energetic electrons (SEEs), namely, which electrons are accelerated in the solar flare and which ones in a CME-driven shock. Solar flares are quick releases of energy and particles from the surface of the Sun, that can be seen as bright flashes of light. CMEs are large ejections of particles and magnetic flux from the Sun. They are often associated with solar flares and sometimes drive shock waves.

Results of the thesis find confirmation of the so-called big flare syndrome – the observation that intense flares are associated with bigger events where phenomena such as fast CMEs and radio bursts appear more frequently. Results of the analysis point at the importance of shocks in producing the most energetic electron events. Findings with elemental abundances suggest that they are not reliable indicators of whether one is observing only flare accelerated particles. Further, this finding hints in the direction of re-acceleration of particles in shocks, as opposed to events where a shock is the only source of acceleration. Findings with spectral indices of the events are in agreement with CME-driven shocks accelerating electrons to flatter spectra. Results also include a statistically significantly harder spectral index for those events with an associated type II burst compared to those without one, with a cut-off value of -4.

Keywords: solar radio bursts, solar energetic electrons, solar flares, coronal mass ejections

Contents

Introduction

The Sun is an active star. Its level of activity increases and decreases in cycles of 11 years – a phenomenon rooted in the magnetic field of the star. The Sun's activity manifests, among other phenomena, as solar flares and coronal mass ejections (CMEs) in the solar atmosphere. Solar flares are short, bright flashes where radiation across the electromagnetic spectrum is emitted. They are the product of magnetic reconnection, that is, a change in the configuration of the Sun's magnetic field that results in a release of energy [1]. These flares are sometimes accompanied by coronal mass ejections that are massive releases of plasma and magnetic flux from the Sun. The plasma that makes up the CMEs can travel at a sufficient velocity in the solar corona and the interplanetary medium to drive a shock wave at its leading edge or even at its flanks.

The magnetic field of the Sun is carried out into the interplanetary medium by the solar wind, an always present stream of particles from the Sun. Because of the rotation of the star, the magnetic field takes on a spiral shape, along which energetic charged particles travel [2]. Standing out from this background solar wind are, for example, the increased intensities of particles accelerated in a CME driven shock, or a solar flare. This kind of increase in particle intensities as seen by an observer is called a solar energetic particle (SEP) event. This thesis focuses on solar energetic electron (SEE) events as observed by the spacecraft Solar Orbiter [3]. In combination with the in-situ particles measurements this work analyses remote observations of CMEs and the radio domain in the form of solar radio bursts. These remote observations are linked with, and are often the first observables of, eruptive phenomena on the surface of the Sun. Solar radio bursts are bursts of electromagnetic emission in the radio domain that stand out from the background radio emission. They are generated by energetic electrons moving through and interacting with the solar atmospheric plasma. There are multiple types of solar radio bursts, classified

by their appearance on a dynamic spectrogram, and some of them can be directly linked to other phenomena on the Sun. Type II radio bursts are related to shock waves, and type III bursts are associated with solar flares, for example.

Two quantities, the maximum energy and intensity of particles in an SEP event, can be thought of as measures of the overall strength of the event observed, and are important not only for spacecraft but for ground-based equipment as well. For the most part, particles of a broad range of energies are harmful to satellites in orbit, and very fast CMEs can drive geomagnetic storms that have the potential to harm electrical infrastructure.

Solar flares and CME-driven shock waves are both potential accelerators of SEEs. The main focus area of this thesis is the determination of the mechanisms and source of energetic electron acceleration from in-situ SEE measurements and remote observations of CMEs and solar radio bursts.

The main unsolved question that this thesis tries to shed light on is the acceleration mechanisms/region of SEEs. The two proponents that are looked at in this study are flare acceleration and CME-driven shock acceleration of SEEs.

1 Theory

The Sun's magnetic field is the source of its activity. It is a complicated structure of interaction between the magnetic field and the plasma that makes up the star. The Sun rotates differentially; it rotates faster at the equator than it does at the poles [4]. As it does so, the magnetic fields get more and more mixed up, more so on the equator than around the poles. One can imagine this is like twisting a rope tighter, storing up energy in the rope. This energy is bound to be released somehow, and what eventually happens is the tangled magnetic field lines reconnect, releasing the energy in the form of solar flares and other signatures of the Sun's activity [5]. The Sun's activity oscillates in periods of around 11 year, during which the star's

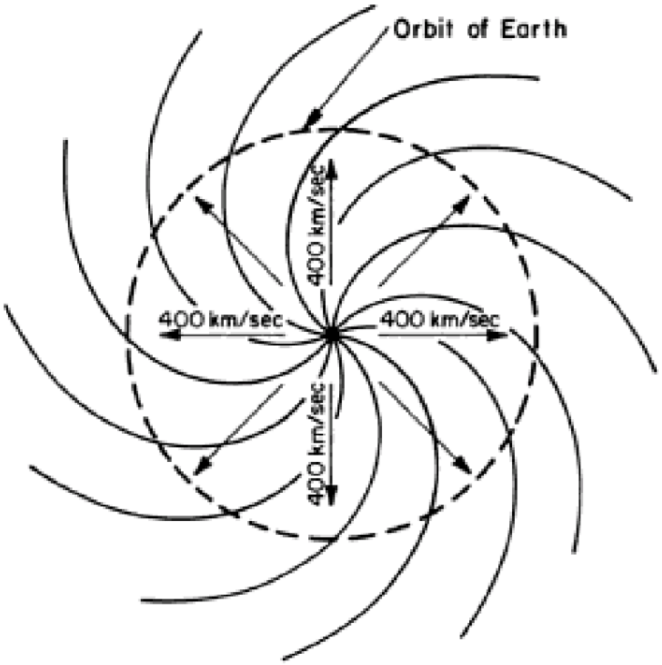


Figure 1: Cartoon of the Parker spirals. The Sun is in the middle, with the view from on top of the north pole [6, 7].

activity goes from one period of minimum activity to another through a period of maximum activity, and at the end of which the solar magnetic field flips in polarity.

The plasma beta β is a quantity of plasma defined by the ratio of the plasma's gas pressure to the magnetic pressure. It describes which of the two forces dominates in the plasma. When $\beta > 1$, gas pressure dominates and any streaming plasma will "drag" the magnetic field along. Similarly, when $\beta < 1$, magnetic pressure dominates and the movement of the plasma is dominated by the magnetic field. β varies greatly with altitude in the solar plasma, but from ≈ 2 solar radii onwards the value becomes ≥ 1; the magnetic field is frozen in into the solar wind and gets dragged along with it [8]. This together with the Sun's rotation makes the magnetic field take on a spiral shape called the Parker spiral, as depicted in Figure 1. The curvature of these spirals depends on the rotation speed of the Sun (which, for the purposes of this work, is a constant at a given latitude) and mainly the speed of the

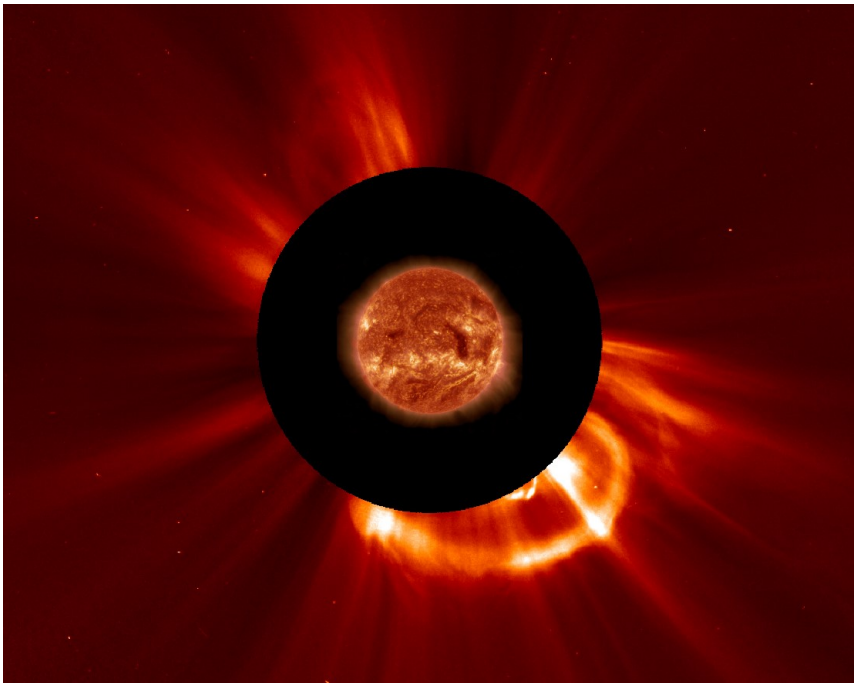


Figure 2: Coronal mass ejection imaged on August 20, 2013 by SOHO LASCO C2 coronograph. It is the bright feature on the south side of the Sun. The Sun is occulted by the black disk seen in the image, with an extreme ultraviolet image of the Sun superimposed on top of it [9]

solar wind (the nominal value usually being 400 km/s). Charged particles emitted from the Sun travel in helix shaped paths along these magnetic field lines.

Textbook SEP events have two phases: the initial rise to peak of particle intensity, followed by a decay phase in which the intensity returns back to background level. SEP events can be crudely categorised into two categories: gradual and impulsive events. Impulsive events are characterised by a sharp rise to peak and a relatively fast decay phase compared to the decay phases of gradual events. They are usually associated with solar flares. Gradual events can have a slower rise to peak and usually a slow decay time that can last for days. Gradual events are usually associated with shocks [2].

To see the particles of an SEP event in-situ, the spacecraft of course needs to be situated on the trajectory of those particles. Since the charged particles travel

along the magnetic field, the spacecraft would then need to be more or less on the magnetic field line that connects it to the source region of the particles. If this is the case, the spacecraft is magnetically well connected to the source region, and the event is deemed *well-connected*.

The energy of an electron in astrophysics most often given in the units of electron volts (eV). The maximum energy of a given SEE is the maximum energy of electrons that is still seen above the background level. Values for the maximum energy have a range of several orders of magnitude with typical values in this thesis being in the 0.1 – 1 MeV range. Particle intensity is a measure of the number of particles seen in a given event by an observer, and has units of $(\mathrm{cm}^2\mathrm{srMeVs})^{-1}$. In an event there are particles of many different energies and the particle intensity can be measured for each energy separately, of course depending on the measuring instrument. The maximum particle intensity for any given energy is the maximum value of the intensity for particles of that energy.

While electrons can be observed in-situ by a spacecraft, certain population of electrons can also be observed remotely via solar radio bursts. Solar radio bursts are bursts of radio emission, produced by energetic electrons. There are different types of radio bursts and some of them are associated with eruptive phenomena from the Sun. This is discussed more in detail in Section 1.4.

1.1 Solar flares

Solar flares, bright flashes of released energy from the Sun, are in many cases the first phenomenon to be observed when studying SEP events. Flares might then be accompanied by CMEs and shock waves, as was the case with the flare imaged in Figure 3. This flare was categorised as X5.4, meaning that an X-ray flux of $5.4 \cdot 10^{-4}\mathrm{W/m}^2$ was measured in the 0.1 – 0.8 nm passband by the NOAA GOES spacecraft [12].

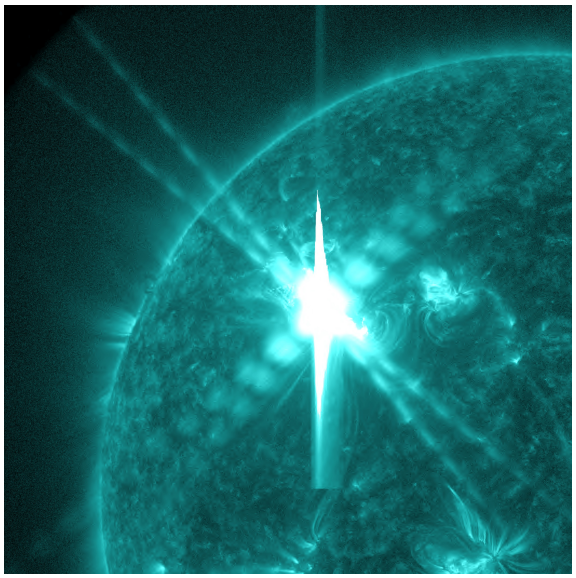


Figure 3: Solar flare observed on March 6, 2012. This flare was also accompanied by a big CME [10].

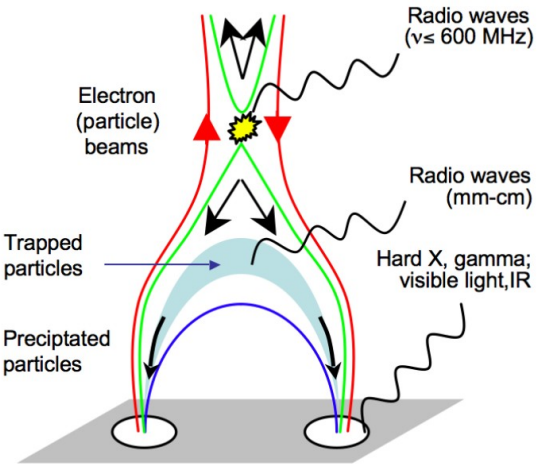


Figure 4: Standard flare model cartoon showing the different particle populations and related electromagnetic emission [11].

Solar flares are sometimes accompanied by CMEs, but a flare can also happen on its own, without an associated CME. A flare that does not have a CME is called confined, and a flare with a CME is called eruptive [13].

Some of the energy released in a solar flares gets converted into kinetic energy of charged particles, thus the flare is a source of SEPs. A cartoon showing the particle populations and related electromagnetic emission is shown in Figure 4. The charged particles released are thought to be the same ones that produce the hard X-ray emission during the flare. The beams of electrons escaping on open magnetic field lines can produce the type III radio bursts.

1.2 Shock waves in the solar corona

A shock wave is a sudden change from supersonic to subsonic flow of a material, when the flow speeds are measured in the frame co-moving with the shock front. The supersonic flow side is called the upstream and the subsonic flow side the downstream. A shock wave forms when some disturbance in the medium moves at a faster speed than the speed at which information about the disturbance can reach the rest of the medium [14]. This information propagation speed in the Earth's atmosphere is called the sound speed. In the solar corona, however, charged particles form a plasma in which the electromagnetic forces between the particles have to be taken into account. In a plasma, the relevant propagation speed is the magnetosonic speed c_{ms} [15].

A CME released by the Sun can generate a shock wave if the propagation speed is sufficient. As the CME propagates, with sufficient speed, a shock wave forms beyond its leading edge to "inform" the incoming material of the moving disturbance. This type of shock wave is called a piston-driven shock, that is, it is not caused by a sudden release of energy like that of an explosion but rather the CME acts as a "piston" that moves fast enough, and drives the shock wave. The solar wind speed

has a nominal value of 400 km/s and the radial propagation speed of CMEs range from a few hundred up to a few thousand km/s.

For the purposes of this study, the most important property of a shock wave is its ability to accelerate electrons – this is discussed further in Section 1.3. The type II solar radio burst (discussed in Section 1.4.1) is a radio indication of the presence of a shock wave in the solar corona.

1.3 Solar energetic particle acceleration

A long-standing problem in solar physics is understanding the acceleration mechanisms/regions of SEEs. There are multiple potential processes, two of which are discussed:

1) Acceleration in a solar flare. The acceleration of charged particles as a result of the solar flare through, e.g., magnetic reconnection driven mechanism [16, 17].

2) Acceleration in shock waves. Two mechanisms (out of several) by which acceleration in a shock wave could happen: diffusive shock acceleration (DSA) and the generally more favoured shock drift acceleration (SDA). The favouritism of SDA is because it accelerates particles fast, and does not, unlike DSA, rely on the particles scattering off magnetic fluctuations.

Particle trajectories at the shock when the particle crosses over from upstream to downstream, and when the particle is reflected in the SDA case are depicted in Figure 5. Whether the particle is reflected or not is dependent on the particle's pitch angle, that is, the angle between the particle's momentum vector and the magnetic field vector. The reason for the drift is that the particle feels the gradient of the compressed magnetic field as it crosses the shock [18].

In the DSA case, acceleration of an electron in a shock wave is essentially a set of frame transformations as the particle crosses over from downstream to upstream, or from upstream to downstream, of the shock, each time gaining momentum. This

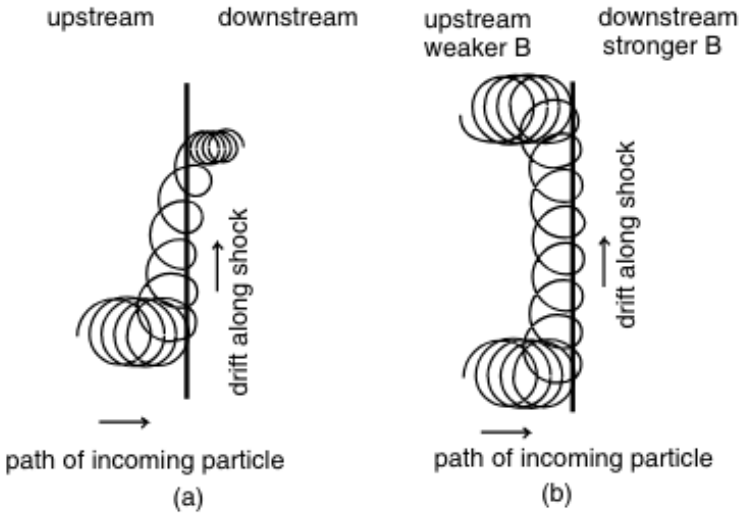


Figure 5: Shock drift acceleration cartoons of trajectories of a charged particle that a) crosses a shock and b) is reflected at a shock [18].

is illustrated in Figure 6 where a trajectory of a particle crossing over a shock is presented. The particle repeatedly diffuses via magnetic turbulence across the shock, which leads to a gain of energy for the particle in the rest frame of the shock [19]. Electrons resonate with high-frequency magnetic fluctuations which are not very intense around shocks, and thus DSA is not able to confine particles to the vicinity of shocks.

In an ideal case, what results from a shock acting on a seed population of electrons is an energy spectrum that follows a power law, where the particle intensity has the following relation:

$$I \propto E^{\gamma}, \tag{1}$$

where E is the particle energy and γ is the so-called spectral index which determines the slope of the power law in a log-log representation.

Another scenario is one in which the spectrum develops a break, or a more gradual roll-over, at some energy. The spectral break is mathematically expressed

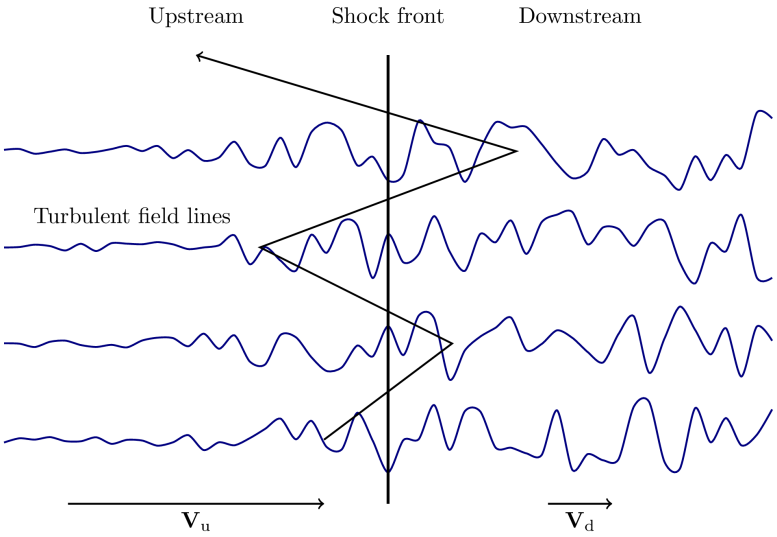


Figure 6: Diffusive shock acceleration cartoon. The black arrow represents a particle crossing over the shock multiple times, on average each time gaining energy [20].

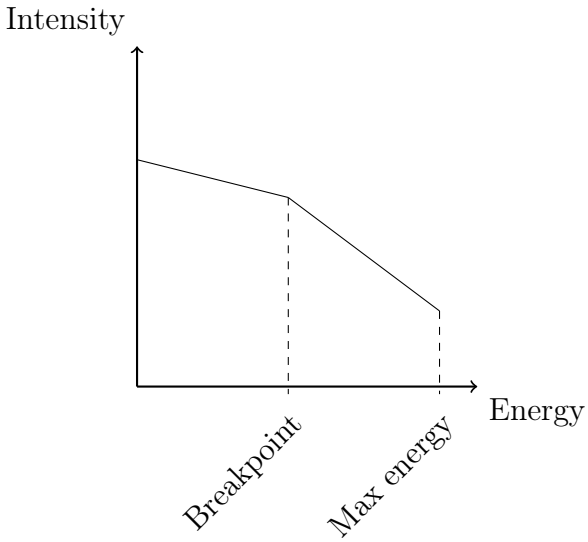


Figure 7: Cartoon of a broken power law particle spectrum, as plotted on a log-log scale. The two different sloped parts of the spectrum are characterised by their own spectral indices.

as a broken power law (depicted in Figure 7) where the two slopes are characterised by their own spectral indices [2, 17].

1.4 Solar radio bursts

The Sun is visible across the electromagnetic spectrum, including radio frequencies. Solar radio bursts refer to features of heightened intensity in the radio domain that stand out from the background radiation. These radio bursts are classified into different types, denoted by Roman numerals. The classification scheme is strictly morphological, that is, it depends only on the distinct shapes of the bursts as seen in a dynamic spectrum plot. In most cases the absolute intensity of the burst is of no interest.

The bursts are produced by beams of up to tens of keV electrons interacting with waves in the plasma they are moving through, resulting in radiation in the form of the so-called plasma emission: Consider a box of plasma consisting mostly of protons and electrons, where the positions of the population of electrons is offset from the population of protons by some arbitrary length. Electrons are much lighter than protons and so the protons would effectively stay still while the electrons would be pulled towards equilibrium by Coulomb forces, eventually leading them to oscillate about the equilibrium positions. These oscillations are called Langmuir waves. Plasma emission is the conversion of energy from the Langmuir waves to electromagnetic radiation as a result of a particle-wave interaction [21, 22]. The frequency at which these plasma oscillations occur is given by the formula

$$f_{\mathrm{pe}}[\mathrm{Hz}] = \frac{1}{2\pi}\sqrt{\frac{n_{\mathrm{e}}e^2}{\epsilon_0 m_{\mathrm{e}}}} \tag{2}$$

$$\approx 9\sqrt{n_{\mathrm{e}}[\mathrm{m}^{-3}]} \tag{3}$$

where n_{e} is the electron density of the surrounding plasma, e the elementary charge,

ϵ_0 the vacuum permittivity and m_e the mass of an electron. f_{pe} is called the plasma frequency and it is important to note that it is proportional to the square root of electron density. Only radiation of higher frequency than the local plasma frequency can propagate in the solar plasma – this means that, given a model of the coronal electron density, the radio emission gives us a lower limit to the height of the emitting electron population.

In this thesis, three types of solar radio bursts are considered: type II, type III and type IV. This is because these bursts are associated with solar eruptive phenomena; type IIs with shock waves, type IIIs with flares and type IVs possibly with CMEs.

1.4.1 Type II radio bursts

Type II radio bursts look like drifting lanes of emission on a spectrogram. The lower frequency lane is called the fundamental which is often, but not always, accompanied by a second lane called the harmonic. The harmonic lane's frequency is twice that of the fundamental's. Secondary lanes of emission can also be present for both the fundamental and harmonic. These different lanes possibly correspond to different regions of emission in and/or around the shock.

Figure 8 has in it a typical type II radio burst with both the fundamental and harmonic frequencies showing. Type IIs are distinguished by their slow drift rate in frequency, which is the result of the emitting source moving at a relatively slow pace up in the solar corona. This source is the electron population accelerated by the slowly moving shock wave, as discussed in Section 1.2. As such, type II bursts are an indication of the presence of a shock wave and most likely a CME that is driving it [24, 25]. The drift rate also implies the speed of the shock wave as it travels through the plasma, which can be recovered by fitting a density model to the burst – methodology that is discussed in Section 3.1.

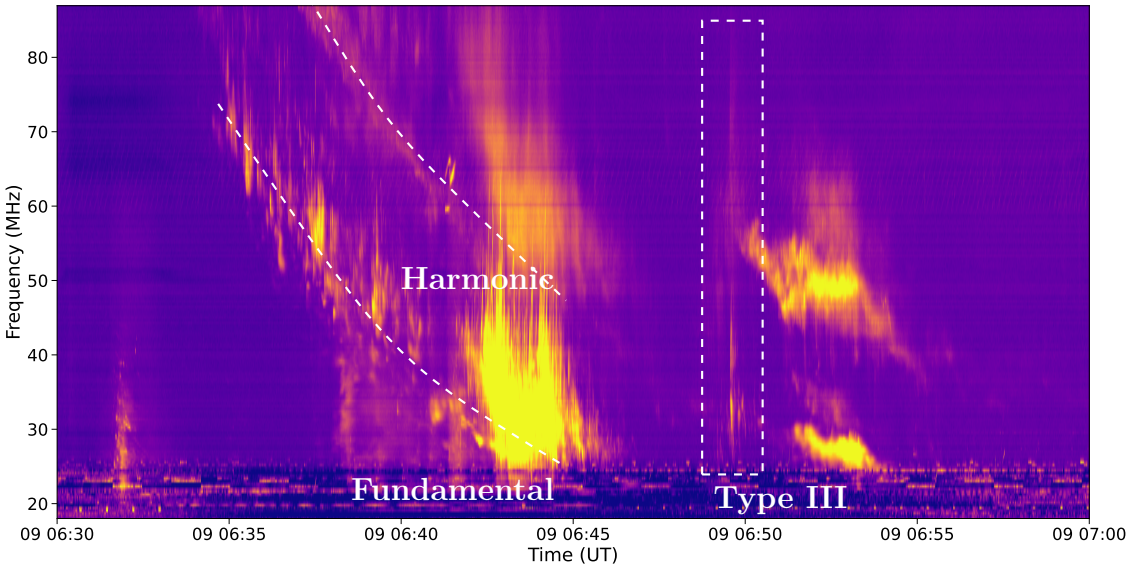


Figure 8: Example of a radio dynamic spectrum showing a type II and type III radio burst. The curved dashed lines roughly show the type II fundamental and harmonic frequency drifts. The dashed box is a type III burst with its much faster drift rate. Data from e-Callisto, observed by Australia-ASSA on the 9th of October, 2021 [23].

1.4.2 Type III radio bursts

Figure 8 also has in it a typical type III radio burst, which can be seen inside the dashed box. The mechanism responsible for the type III is the same as the type II but from the drift rate difference it is obvious that emitting sources producing type IIIs move much faster than sources that produce type IIs. The speed of these electrons is typically around 0.1 to 0.5 times the speed of light ($\approx 3 - 80$ keV). A well established idea is that type IIIs are produced by beams of electrons moving along open magnetic field lines. A classical idea is that these electrons are accelerated by magnetic reconnection in reconnection sites such as flares [17]. It should also be noted that nowadays, with much higher resolution instruments, type IIIs appear to be present even in the absence of large eruptive phenomena [26].

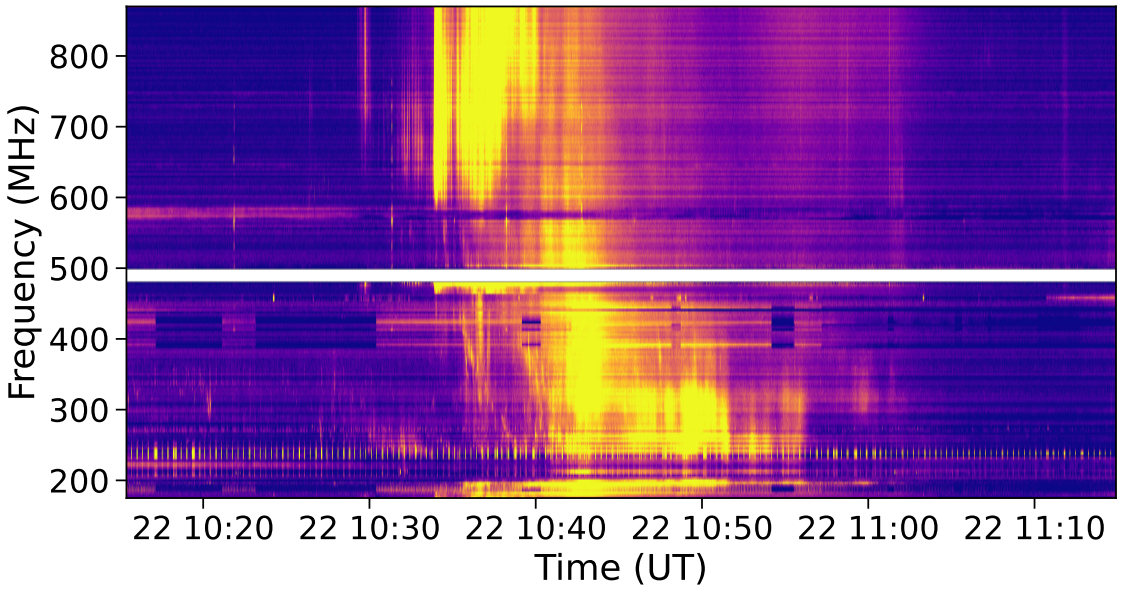


Figure 9: Examples of type IV radio bursts at 200–400 MHz, and at above 600 MHz. Observed by e-Callisto BLEIEN on September 22nd 2011.

1.4.3 Type IV radio bursts

Type IV bursts are elusive. Literature often cites them as broad-band continuum emission [27] – an example can be seen in Figure 9 – which can be either moving or stationary, potentially attributed to some emitting population of electrons trapped in, say, the loops of CMEs or the post-eruption loops behind a CME. This is supported by the fact that the moving type IVs are for the most part only seen in the presence of a CME [27, 28].

1.5 Elemental abundances

The solar photosphere, corona as well as particles seen in SEP events all have some typical values for the composition of elements and their ions [29]. A peculiar finding is the enhanced ^{3}He/^{4}He and Fe/O ratios that are often associated with impulsive SEP events. This might suggest that measuring certain enhanced isotropic abundance ratios, like ^{3}He/^{4}He, would mean that one is looking at the flare produced

component of that event. The enhanced isotropic abundances are not usually associated with gradual SEP events [30, 31].

2 Data

The data selected for this analysis consist of in-situ particle measurements, radio data of solar radio bursts and white-light data of CMEs, and are discussed in more detail in this section. All quantities, their symbols, units and descriptions can be found in Appendix A. The Solar Orbiter dataset is described in Appendix B.

2.1 Energetic particle data

Electron data are the focus of this study. The main source of data is the Solar Orbiter mission from a time frame beginning in 05/2021 and ending in 11/2022. The data are in the form of an event list compiled by the Solar Orbiter multi-instrument working group*. The in-situ particle measurement suite on-board Solar Orbiter [3] is called the Energetic Particle Detector (EPD) [32], in which the three instruments that are capable of measuring electrons are the SupraThermal Electrons and Protons (STEP), the Electron Proton Telescope (EPT) and the High Energy Telescope (HET). The main observable in the particle data is the peak particle intensity, which in this study is measured in units of $(\mathrm{cm^2 srMeVs})^{-1}$ at an energy of 43 keV which is the energy used in the event list unless specified otherwise. Together the three instruments are capable of measuring a good portion of the energy spectrum from which quantities such as the maximum electron energy and spectral indices of the events are derived. The determination of the spectral indices as well as the break point energy was done by fitting an appropriate power law model to the spectra. The maximum energy for a given event is defined as the middle of

*Warmuth *et al.* (in preparation)

the highest energy bin that is significantly above the background noise. Spectral analysis was done by Annamaria Fedeli*.

The data from Solar Orbiter are compiled originally in a list of events with the focus on events where Solar Orbiter observes both the solar flare and the corresponding SEE event. Part of the selection criteria is that a temporal association between the flares and SEE events could be made. For this thesis, from this set of events was then picked ones that could be "seen" on Earth, meaning that the longitudinal separation of the flare site was required to be less than 100 degrees as seen from Earth. 100 degrees separation, as opposed to say 90 degrees, was chosen because radio waves can be seen from slightly behind the limbs of the Sun. This was done to create a comparable set of events for which ground-based instruments could detect radio bursts. The motivation for using radio data in combination with the information of the event list is to determine if a shock is also present in each of the events.

The Solar Orbiter event list contains the X-ray intensities of the flares observed. These are not measured by Solar Orbiter but by the X-ray Sensor instruments onboard the GOES satellites [12, 33]. The data containing the Solar Orbiter measurements as well as the GOES flare intensities will hereafter be referred to as simply the Solar Orbiter dataset.

This thesis also analyses a bigger set of electron data from the Wind spacecraft [34], spanning a time range of around 07/1996 – 11/2017 and including over 1500 events†. The Wind list only has information about the time and minimum and maximum energy of the events, so it is not as comprehensive as the Solar Orbiter data. The data are in a form of an event list, compiled from data from the 3DP instrument [35]. GOES X-ray intensities are not present in this set of data.

*Fedeli *et al.* (in preparation)

†The Wind event list was obtained from Dr. Linghua Wang, private communication.

2.1.1 Abundance ratio

Included in the EPD suite is also the Suprathermal Ion Spectrograph (SIS) whose elemental composition data, provided in the list of events, is used in this study.

The two areas of interest are the ^{3}He/^{4}He and the Fe/O ratios and they are coded as simple binary variables (no, yes) in this thesis – a 'yes' in ^{3}He/^{4}He meaning that ^{3}He ions are seen in the SIS data, and a 'yes' in Fe/O meaning a ratio of around unity. These definitions come from list of events that was used.

2.2 Radio data

The radio data used are from National Oceanic and Atmospheric Administration (NOAA) [36], observed by the Radio Solar Telescope Network (RSTN). RSTN is a set of 7 radio instruments located in Culgoora, Australia; Holloman Air Force Base, New Mexico, USA; Palahua, Hawaii, USA; Sagamore Hill, Massachusetts, USA; Learmonth, Australia; Ramey Air Force Base, Puerto Rico, USA; and San Vito, Italy.

The ionosphere of Earth becomes quite opaque in the tens of megahertz and because of this the RSTN network observes the Sun in the metric range. This is also the frequency range in which solar radio bursts occur. Observations of type II, III and IV radio bursts are used in this thesis because of their relation to the solar eruptive phenomena, which ties in with the in-site particle observations.

The radio data used come mainly in the form of binary (no, yes) variables representing the presence of a certain type of solar radio burst. The association to a particular SEP event is purely temporal in this analysis, meaning a burst was deemed associated with an event if it was observed around the time of the event. Type II bursts are of particular interest because they carry information about the presence of and the propagation speed of the shock wave that accelerates the radio burst emitting electron population. Because of this, also the speeds of the shock

waves inferred from the type II bursts are considered.

2.3 Coronal mass ejection data

The presence and propagation speed of CMEs were considered in the study. The propagation speed is the first order linear approximation, estimated from white-light imaging. The presence of one is coded in a binary variable. CME data were obtained from the Large Angle and Spectrometric COronagraph (LASCO) catalogue [37]. LASCO is a set of three (one is not operational, so two were used in this thesis) coronographs on board the Solar and Heliospheric Observatory (SOHO) spacecraft, situated at the Lagrange point 1 in between the Sun and Earth [38]. It produces white-light images, like the one in Figure 2, which can be used to detect CMEs erupting from the Sun.

3 Methodology

The set of events seen by Solar Orbiter were selected with the criterion that the separation angle between the flare site in the event list and the Earth in the plane of the solar system is less than 100 degrees, as depicted in Figure 10. Since only events where Solar Orbiter measured significant particle intensities on were selected, the analysis mainly deals with SEE events that are well-connected to the source region. Thus no further methodology for checking the magnetic connection is needed.

One of the basic methodological ideas of this analysis is looking at separately those events that have type II bursts (as observed by ground-based equipment) or CMEs associated with them, and those that do not. This is reflected in many of the plots presented.

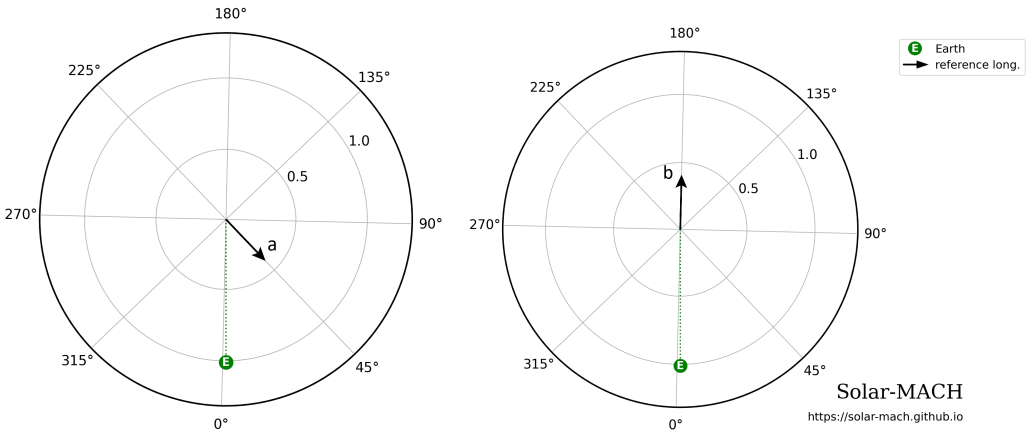


Figure 10: Depiction of the angle separation selection criterion. In the two images the Sun is in the middle and the green circle is Earth, viewed from above the solar system plane. The arrows, denoted **a** and **b**, represent solar flares. Flare a passes the selection criterion because it is less than 100 degrees separated from Earth, whereas flare b does not as it is more than 100 degrees separated [39].

3.1 Measuring shock speeds

To determine shock speeds from type II radio bursts, a Python tool was written. The basic idea is that for each (time, frequency) point in a dynamic spectrogram (that is part of the type II), one can determine the height of the shock from the frequency via Equation 2 and a density model. The density model chosen here is the common Newkirk model [40] that has the form

$$n_{\mathrm{e}}[\mathrm{cm}^{-3}] = 4.2 \cdot 10^{4.0+4.32/R}, \tag{4}$$

where R is the distance from the center of the Sun in solar radii. Equation (4) together with the plasma frequency Equation (2) can be used to derive the following equations:

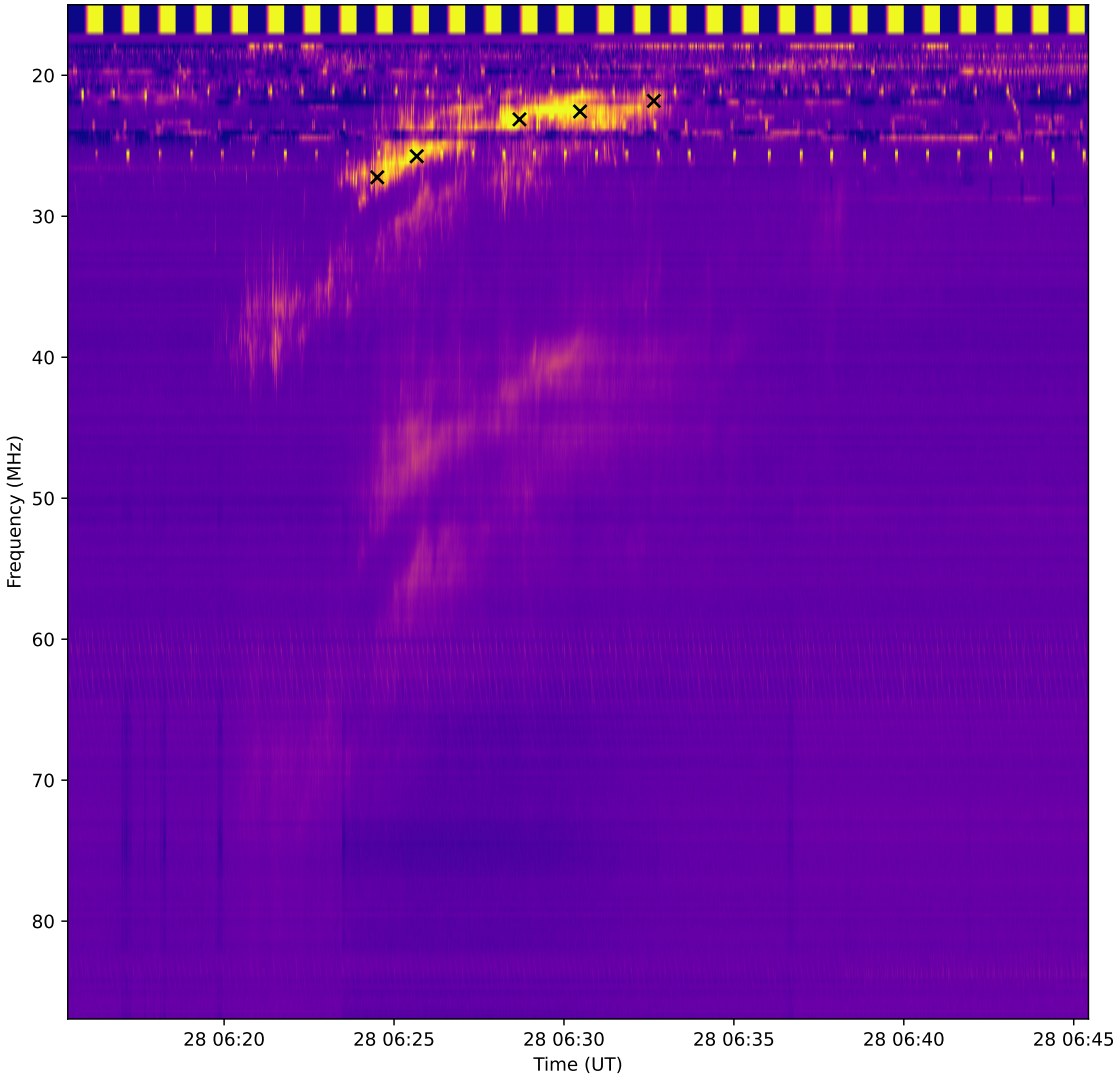


Figure 11: Points (the black x's) clicked on a dynamic spectrum of a type II radio burst. For each point the height of the shock is determined via the density model.

$$h[\mathrm{R}_\odot] = \frac{4.32}{\log 10(n_\mathrm{e}/a \cdot 4.2 \cdot 10^4)}, \tag{5}$$

$$n_\mathrm{e}[\mathrm{cm}^{-3}] = \left(\frac{f \cdot 10^6}{8980 \cdot n}\right)^2, \tag{6}$$

where f is frequency, n is the harmonic number (1 for the fundamental lane of a type II burst) and a is the number of so-called "folds" in the model.

The tool allows for the determination of the height at multiple (time, frequency) points and from this is able to calculate the propagation speed of the shock, like is shown in Figure 11.

The radio data source mentioned in Section 2.2 also includes the shock speeds but this way of determining the speed by hand should be more accurate, because it avoids all the potential flaws that could be present in the automatically determined data. Determination by hand was done for all type IIs in the Solar Orbiter dataset, if radio data of the burst could be acquired from the e-Callisto catalogue. This was not done for the bigger set of Wind data due to the large number of events and lack of automation of the process.

3.2 Permutation test

Suppose we have two sets of observations, A and B. A basic way of testing whether the two sets of observations are sampled from the same or a different distribution is to measure the means of the sets observations $\bar{A}$ and $\bar{B}$. One then usually wants to compare the two means, so for this section the difference of the means $\bar{A} - \bar{B}$ is called the *statistic* that we want to test. This approach works if the distribution from which the observations are measured is normal, and the sample sizes are large. However simply calculating the means doesn't result in any approximation for uncertainties.

A permutation test is a way of resampling data that tests against the null hypothesis that all of the observations from A and B come from the same distribution.

There are different ways in which one can set up the test but the general idea goes as follows; take a subset of observations from A and B, say a and b, and calculate the statistic (e.g. $\bar{a} - \bar{b}$) that you are interested in. Repeat this process for a sufficient number of iterations. The result is a distribution of the values of the statistic. By comparing this distribution, and the value of the original statistic (e.g. $\bar{A} - \bar{B}$) one can compute an approximation for the uncertainty of the statistic.

The permutation tests done in this thesis were implemented using SciPy's permutation_test() function, using 100000 iterations [41].

3.3 Correlation tests

Two of perhaps the most common correlation tests – the Pearson and the Spearman tests – were used. The Pearson correlation coefficient r_p is a single value measure of linear correlation between two sets of data, bounded between values of -1 and 1; 1 representing perfect correlation and -1 representing perfect anti-correlation.

The Spearman rank correlation coefficient r_s is defined as the Pearson correlation coefficient computed on ranked variables of two sets of data. It is also bounded between values of -1 and 1, and it in essence says how well the relationship of the two sets of data can be described by any monotonic function, not just a line as is the case with r_p.

Under the assumption that the two sets of data are independent and normally distributed, the quantity $t = \frac{r}{\sigma_\mathrm{r}}$ (where r is r_p and r_s for the Pearson and Spearman correlations, respectively, and σ_r the sample standard deviation) follows the Student's t-distribution. From the computed value of t and the t-distribution we can determine the p-value, which is a measure of how likely it would be to observe the measured correlation by chance, that is, in a case where there truly is no correlation [42]. A p-value of 0.05 is taken as the upper limit for statistical significance in this study. Both correlation coefficients were calculated using functions provided in the

SciPy library for Python [41].

3.4 Logistic regression

Because the data set in this thesis includes many binary variables, i.e., those that can take on a value out of only two possibilities, and because the interpretation of the correlation coefficients for binary variables is ambiguous, logistic regression was used to determine correlations of binary variables.

Logistic regression is a type of linear regression model that is used when the response variable is binary in nature. The model produces S-shaped curves $p(x)$ that are defined (for the case of one predictor x and one response variable) by the equation

$$p(x) = \frac{\mathrm{e}^{\beta_0+\beta_1 x}}{1+\mathrm{e}^{\beta_0+\beta_1 x}} \tag{7}$$

where β_0 is called the intercept term and β_1 the predictor term, and look like ones depicted in Figure 12. Although the terms in Equation (7) are somewhat similar to the intercept and predictor terms when fitting a straight line, the interpretation is not. The intercept term moves the curve along the axis of the predictor variable and hence its main purpose is to simply take on an appropriate value for the scale of the predictor values. The interpretation of it is the log of the odds (log odds) of the event being predicted by $p(x)$ when $x = 0$. This is typically – like in this thesis – not of much interest because of the different implications of $x = 0$ in different physical situations. A value of $p(x) = 0.5$ is achieved at $x = -\beta_0/\beta_1$.

The predictor term β_1 is related to the actual correlation between the predictor and the response, and is thus of more importance. The slope term of a straight line measures how much the response changes when the predictor changes, and this is true for every value of the predictor because of the linearity of a line. The S-shaped curves of equation (7) are not linear, meaning that the change in the p also

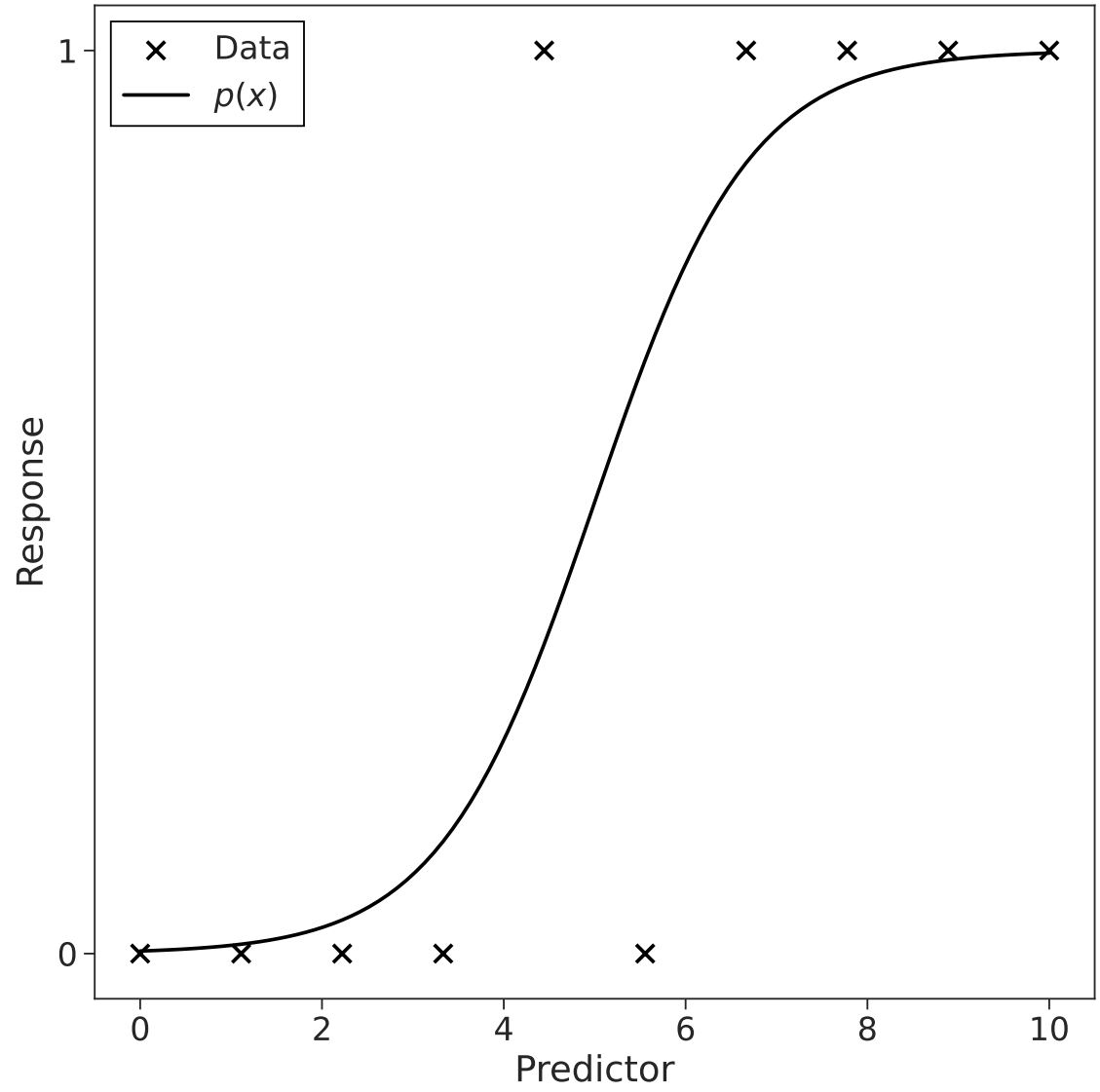


Figure 12: Logistic regression fit. Depiction of the S-shaped curves produced by Equation (7).

depends on the value of the predictor variable. The interpretation of β_1 is that a one unit change in the predictor variable changes the log odds by a factor of β_1, or equivalently the odds change by a factor of e^{β_1}. $1/\beta_1$ also scales with the width of the region where $p(x)$ changes from close to 0 to close to 1. Say that we get an odds value of 2, this means that a unit increase of the predictor variable changes the "likelihood" of seeing the response by a factor of 2. Similarly, given an odds value of 0.2, say, would mean that a unit increase of the predictor variable changes the likelihood of seeing the response by a factor of 0.2, thus decreasing the likelihood and making the correlation negative. Thus the interpretation of the sign of β_1 is the same as in fitting a line: a positive β_1 implies positive correlation, a negative β_1 implies anti-correlation.

The equivalent of the t-distribution for the case of the parameters in logistic regression is the z-distribution. From this distribution one can calculate the p-values of the parameters in the same fashion as from the t-distribution. Just like with the correlation tests, the p-value of 0.05 is taken as the upper limit for statistical significance for the logistic regression parameters [42]. The regression was implemented with the statsmodels library for Python [43].

It is also noteworthy to mention that all of the binary variables in these data represent the presence or absence of something e.g. the presence or absence of a type II or a CME. This means that all of the regression results can be interpreted comparatively, because the value 0 means the same thing in both cases. Opposite to this would be for example that 0 and 1 are encoded to mean different colours, and the value of 0 would not have the same meaning for two different variables.

4 Results

The main findings are that the majority of type II radio bursts occur in events that reach the highest maximum electron energies, and that observations of ^{3}He or Fe

Table I: Correlation coefficients of flare intensity, particle intensity and maximum energy.

	Pearson (coef) (p-value)	Spearman (coef) (p-value)
log intensity vs log GOES	0.48 0.0013	0.49 0.0009
log max E vs log GOES	0.66 < 0.0001	0.57 0.0003
log max E vs log intensity	0.73 < 0.0001	0.76 < 0.0001

rich events does not directly correspond with observing only the flare component of the event. Flare intensity, peak electron intensity and maximum electron energy also scale up together statistically in the data presented.

Figure 13 shows three quantities of the Solar Orbiter data – GOES X-ray flare intensity, peak electron intensity at 43 keV and the maximum electron energy reached – plotted against each other. In the three plots, crosses mean that no CME was associated with the event and circle means that a CME was detected. Red colour means no associated type II burst was detected while green colour means a type II was seen. Good correlations between all of the three quantities are obtained as is shown in Table I. The highest correlation was achieved between maximum electron energy and maximum electron intensity with 0.73 and 0.76 as the Pearson and Spearman correlation coefficient, respectively. This means that all of the three vari-

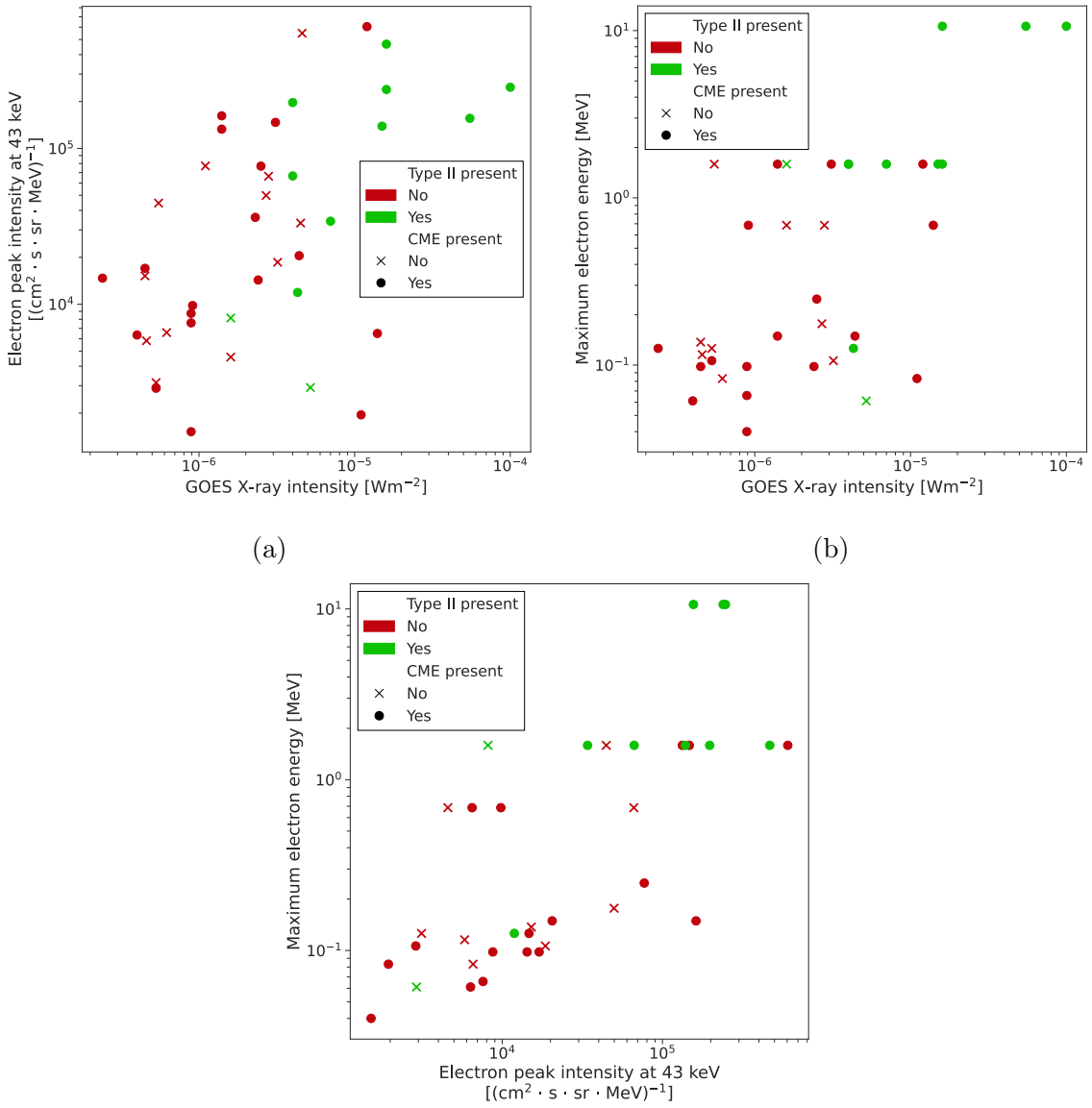
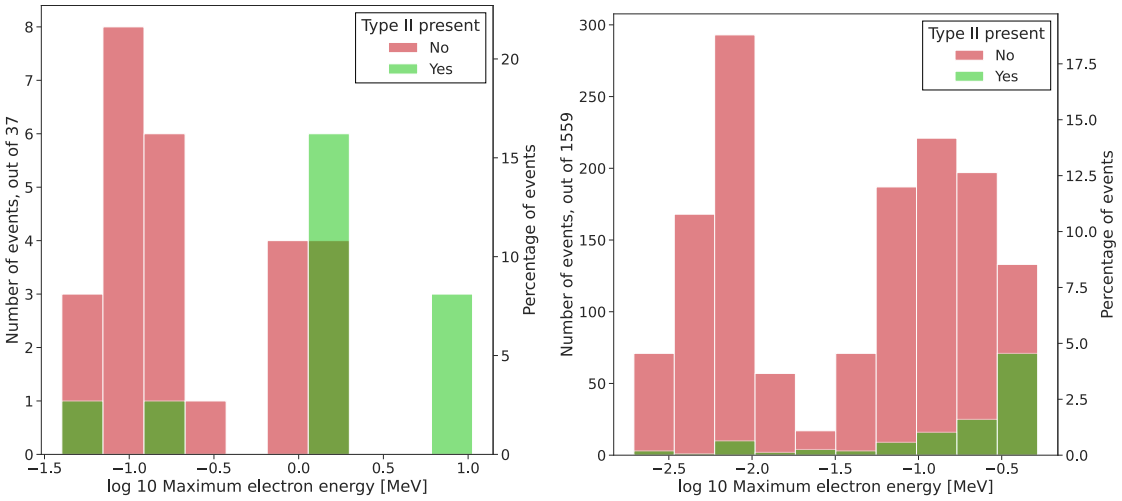


Figure 13: Scatter plots showing a) peak electron intensity at 43 keV vs flare X-ray intensity, b) maximum electron energy reached vs flare X-ray intensity, and c) maximum electron energy reached vs peak intensity at 43 keV. Crosses mean no CME was present, circles mean a CME was present, and the markers are color coded such that red means no type II radio burst and green means a type II was observed.



(a) (b)

Figure 14: Histogram showing the maximum electron energies reached by the events observed by Solar Orbiter (left) and Wind (right). Green means type II burst was observed, red means no type II was present.

ables scale up together statistically. We also notice that the events that have a type II associated with them tend to be in the upper right corners of the plots in Figure 13, meaning they are present in the most energetic and intense events.

Figures 14a and 14b show maximum electron energy histograms of the Solar Orbiter and Wind data, respectively. Both plots contain each event for which radio data was available. The binning is chosen arbitrarily such that there are 10 bins, and the colouring is such that red bins indicate events where no type II burst was detected and green bins those where a type II was detected. We notice that the majority of type IIs appear in the higher maximum energies in the Solar Orbiter data, and this trend gets more pronounced in the Wind data with a bigger sample size. Both histograms appear to have a bimodal distribution; this is most likely an effect due to instrumentation, and is further discussed in Section 5. Scatter plot of the Wind data can be found in Appendix C.

Logistic regression results for the binary variables in the Solar Orbiter data versus

Table II: Logistic regression results of the SolO data for the binary variables against GOES flare intensity. *Coef* is the predictor term in Equation (7). *Odds* is e to the power of *coef*.

	Coef	Std. err.	p-value	Odds
Slow CME vs log GOES	0.8142	0.718	0.257	2.257
Fast CME vs log GOES	10.39	9.920	0.295	32470
Type II vs log GOES	3.364	1.111	0.002	28.90
Type III vs log GOES	-0.8574	0.6080	0.159	0.4243
Type IV vs log GOES	3.449	1.293	0.008	31.48
^{3}He rich vs log GOES	-2.209	0.8120	0.007	0.1098
Fe rich vs log GOES	-0.6748	0.5140	0.189	0.5093

Table III: Logistic regression results of the SolO data for the binary variables against the maximum electron energy. *Coef* is the predictor term in Equation (7). *Odds* is e to the power of *coef*.

	Coef	Std. err.	p-value	Odds
Slow CME vs log max E	0.4120	0.626	0.510	1.510
Fast CME vs log max E	2.193	1.169	0.061	8.965
Type II vs log max E	2.292	0.788	0.004	9.892
Type III vs log max E	0.1661	0.667	0.803	1.181
Type IV vs log max E	37.78	2070	0.985	$2.559 \cdot 10^{16}$
^{3}He rich vs log max E	-0.1699	0.488	0.728	0.844
Fe rich vs log max E	-0.9084	0.619	0.142	0.4031

the logarithm of GOES flare intensity, "log GOES", are given in Table II. The same correlations against the logarithm of maximum electron energy, "log max E", are given in Table III. The important columns are the p-value for the significance of the result and the *odds*. The odds is the factor by which logistic regression function p gets multiplied when increasing the predictor variable by one unit; the bigger the odds, the stronger the correlation of the predictor log GOES and the binary response. In essence this is a measure of how the likelihood of observing, say a type II burst, changes when the flare gets stronger.

In the Solar Orbiter data, presence of type IIs as well as type IVs is statistically significantly correlated with flare intensity. The observation of ^{3}He rich events is statistically significantly anti-correlated with flare intensity in the Solar Orbiter data.

CMEs in Tables II and III are divided into two groups: slow CMEs have a speed lower or equal to 1000 km/s, and fast CMEs are faster than 1000 km/s. The presence of slow CMEs doesn't appear to be statistically significantly correlated with the flare strength.

Table II and III have also the correlations of the observations of enhanced elemental abundances ^{3}He/^{4}He, denoted ^{3}He rich, and Fe/O, denoted Fe rich.

In the regression with fast CMEs in Table II, 42% of the data points were able to be "perfectly" predicted, thus indicating something called complete or quasi-complete separation of the variables. This is essentially when the choice of predictor variable completely or almost completely separates the response variable into groups paradoxically "too well". This is not uncommon with logistic regression, and in this case the cause of this is simply the lack of data; there are only five fast CMEs in the data [44].

Logistic regression results for the Wind data set are given in Table IV. The elemental abundances are not available for this set of data. The presence of fast

Table IV: Logistic regression results of the Wind data for the binary variables against the logarithm of maximum electron energy. *Coef* is the predictor term in Equation (7). *Odds* is e to the power of *coef*.

	Coef	Std. err.	p-value	Odds
Slow CME vs log max E	0.3196	0.082	<0.0001	1.377
Fast CME vs log max E	1.720	0.232	<0.0001	5.584
Type II vs log max E	1.772	0.193	<0.0001	5.884
Type III vs log max E	0.5778	0.074	<0.0001	1.782
Type IV vs log max E	1.590	0.229	<0.0001	4.903

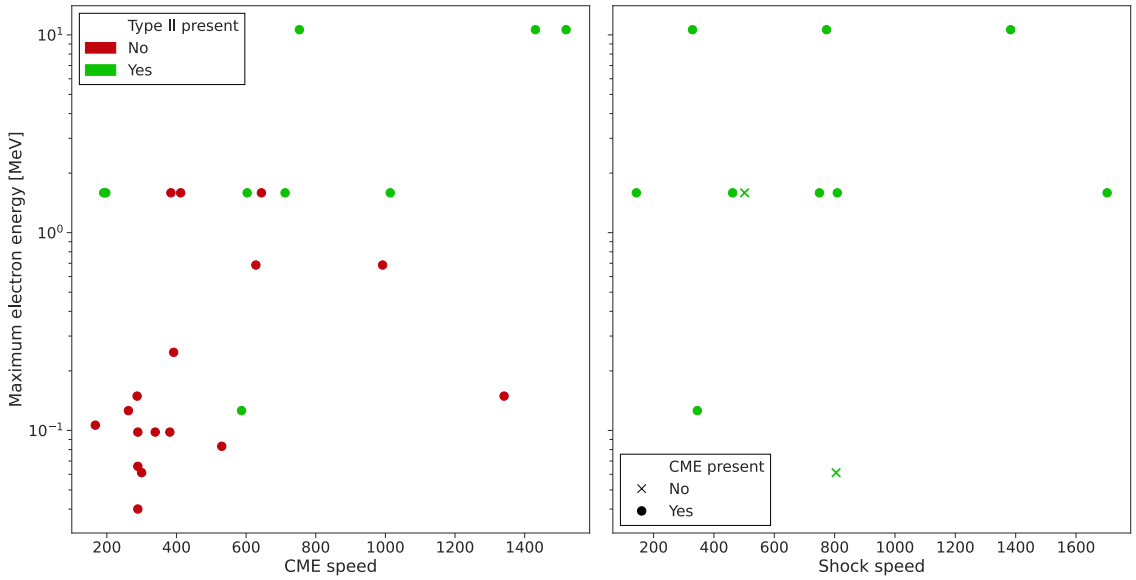


Figure 15: Maximum electron energy as a function of CME speed (left) and shock speed (right) in the Solar Orbiter dataset.

CMEs and type IIs are more strongly correlated with the maximum energy than the presence of slow CMEs and type IIIs; the same pattern that was in the Solar Orbiter data. The correlation of the presence of type IVs appears to be in the middle of these; less correlated than fast CMEs and type IIs but more so than slow CMEs and type IIIs. This suggests the same tendency for the presence of type IVs to increase as the flare gets stronger, as is most likely the case with type IIs and fast CMEs in the big flare syndrome (BFS) [45].

The maximum electron energy as a function of CME and shock speed can be seen in Figures 15 and 16 for the Solar Orbiter and Wind datasets, respectively. The corresponding correlation coefficients are in Tables V and VI. One clear outlier was removed from the Wind dataset due to its very high shock speed and a very low corresponding CME speed. There is no correlation with the shock speed in either dataset. The correlation of around 0.5 for CME speed in the Solar Orbiter dataset is no longer present in the Wind data.

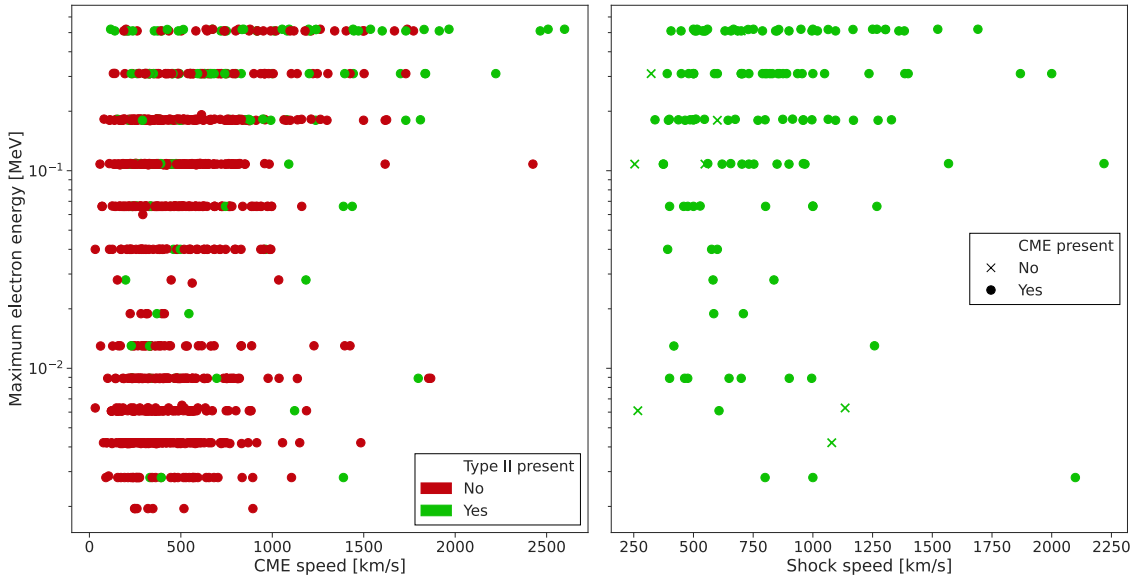


Figure 16: Maximum electron energy as a function of CME speed (left) and shock speed (right) in the Wind dataset.

Table V: Correlation coefficients of the logarithm of maximum electron energy vs. CME and shock speeds, in the SolO dataset.

	Pearson (coef) (p-value)	Spearman (coef) (p-value)
log max E vs CME speed	0.56 0.0029	0.53 0.0054
log max E vs shock speed	0.17 0.62	0.060 0.86

Table VI: Correlation coefficients of the logarithm of maximum electron energy vs. CME and shock speeds, in the Wind dataset.

	Pearson (coef) (p-value)	Spearman (coef) (p-value)
log max E vs CME speed	0.29 <0.0001	0.27 <0.0001
log max E vs shock speed	0.057 0.50	0.17 0.039

Table VII: Permutation test results for the difference of mean spectral index values for two sets of events: those with type II bursts and those without, at 43 keV and at 200 keV separately.

	Difference of means	p-value
43 keV	1.110	0.0195
200 keV	0.4039	0.195

4.1 Spectral indices

Figure 17 shows the spectral index of the electron events at 43 keV (left) and at 200 keV (right), as a function of the flare intensity. There are two spectral indices, because the data is modelled by a double power law i.e. two power laws separated by a break. The 200 keV plot has fewer points on it because many of the events simply do not reach an energy of 200 keV. Two data points are removed from the data, one because of very large errors and one due to unphysical values.

Table VII shows permutation test results for the difference of means of the sets

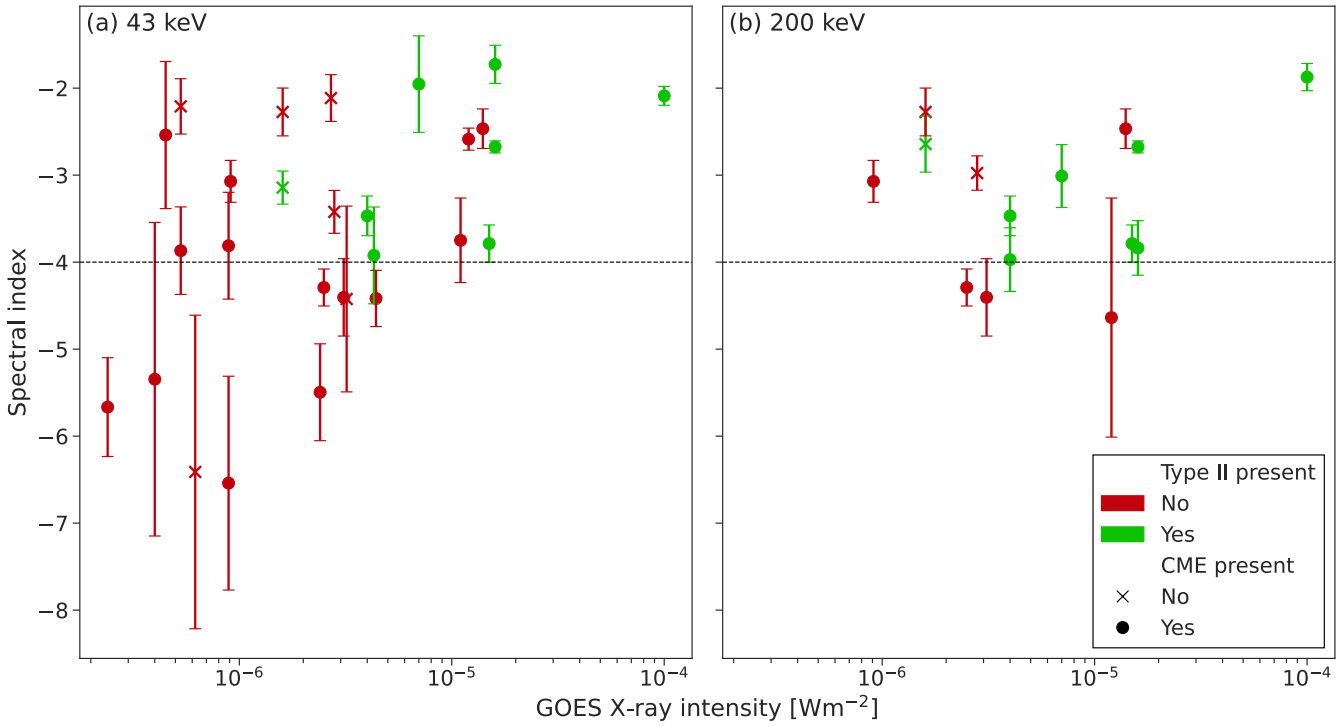


Figure 17: Spectral indices of the events at 43 keV (left) and at 200 keV (right), as a function of the flare intensity.

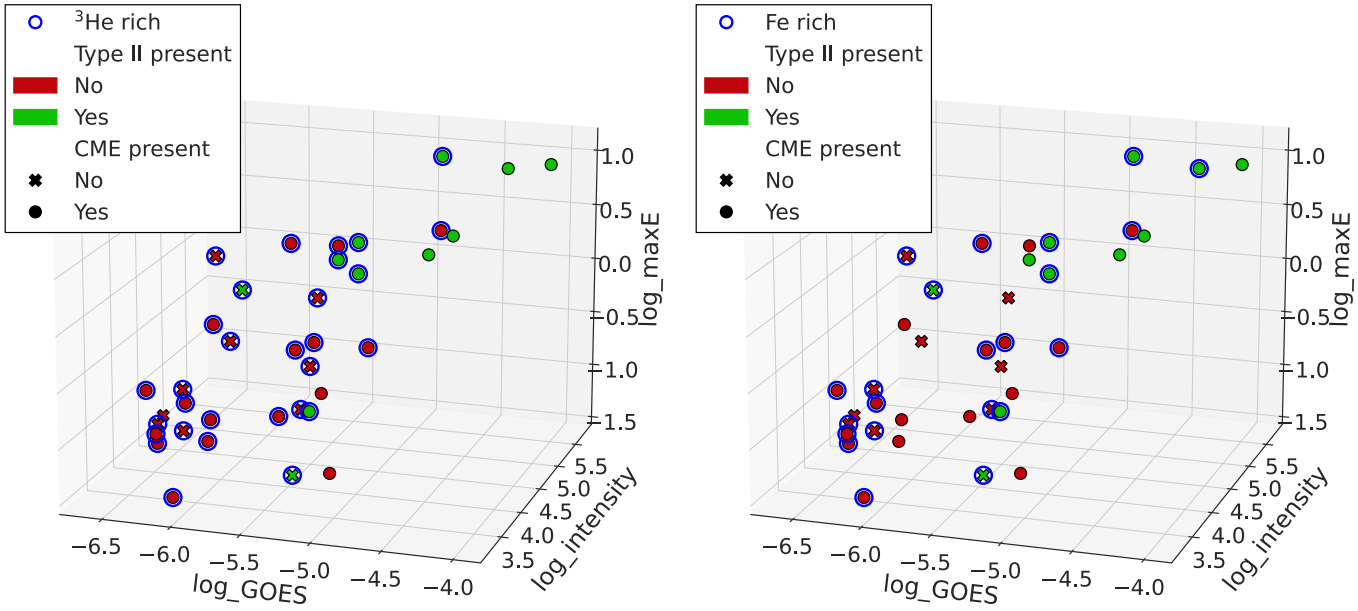
of events that have type IIs associated with them and those that don't, separately for 43 keV and 200 keV. The difference of the means of these two sets of events at 43 keV is around 1.1, and the difference is statistically significant. At 200 keV, the difference is not statistically significant.

We see that at 43 keV, the events that had a type II associated with them show on average harder spectral indices. The trend is no longer as strong at 200 keV. Comparing the two plots it seems more likely that the trend is due to the presence of soft values of spectral indices at 43 keV. This means that the effect of the shocks in the two plots is more or less the same, which is in agreement with the notion that shocks accelerate electrons to flatter spectra [2]. We also see that no events in Figure 17 that have a type II have spectral indices softer than -4, which is marked by the dashed line.

No significant correlation was found in the difference of the spectral indices at 43 keV and 200 keV vs the flare intensity for the events. Good correlations between the spectral indices and CME speeds were also not observed.

4.2 Abundances

Figure 18 shows the Solar Orbiter events in 3D scatter plots. The x- and y-axes are logarithms of the flare and particle intensity, and the z-axis is the logarithm of maximum electron energy. The correlation of all three of these quantities can be seen. Circles correspond to events with an associated CME and crosses to those where a CME was not observed. Green corresponds to events where a type II was observed and red to those where one wasn't observed. The events that are ^{3}He or Fe rich are circle in blue in Figures 18a and 18b, respectively. Knowing these abundances are a result of the acceleration process in the solar flare itself, as was discussed in Section 1.5, one might expect the ^{3}He and Fe rich events to not include indications of CMEs or shocks. This is not what is observed; instead we see the



(a) (b)

Figure 18: 3D scatter plots showing the data on a log10 X-ray intensity vs log10 peak intensity vs log10 maximum energy axis. Blue circled markers indicate a) events that are ^{3}He rich and b) events that are Fe rich.

^{3}He or Fe rich events being scattered seemingly without any regard for the other variables.

Figure 19 shows the radial distance and magnetic footpoint separation of the Solar Orbiter events. The magnetic footpoint separation is the angle difference of the flare site and the connection point of the magnetic field line that Solar Orbiter is on; the closer this value is to zero, the better connected Solar Orbiter is to the flare site, in principle. The radial distance is the line-of-sight distance from Solar Orbiter to the Sun. Figure 20 shows otherwise the same plot but with the presence of type II bursts, instead of CMEs, indicated by black circles. Logistic regression results of the GOES X-ray flare intensity and ^{3}He and Fe rich events are given in Table II. ^{3}He rich events are statistically significantly anti-correlated with flare intensity, however, Fe rich events are not.

5 Discussion

The findings of this thesis hint at shocks being the source of acceleration for most intense and energetic SEE events. They also point in the direction of re-acceleration in shocks of electrons that originated from flares.

Figure 13 and the correlation results of Table I are indicative of the so-called big flare syndrome (BFS) [45]. The BFS means that the biggest flares are associated with fast CMEs and strong shocks, making it more difficult to distinguish where the observed SEEs were accelerated. The flare is not necessarily a cause of any of the other observables but rather that the flare intensity can be thought of as one of the measures of the overall strength and energy release in a given event.

The observation that the majority of type IIs in Figure 13 correspond to the high flare intensity events is also a part of the BFS; the biggest events also have shocks even though the flare is definitely not the cause of them. Perhaps more importantly this tells about the role of the shock in producing the most energetic and intense

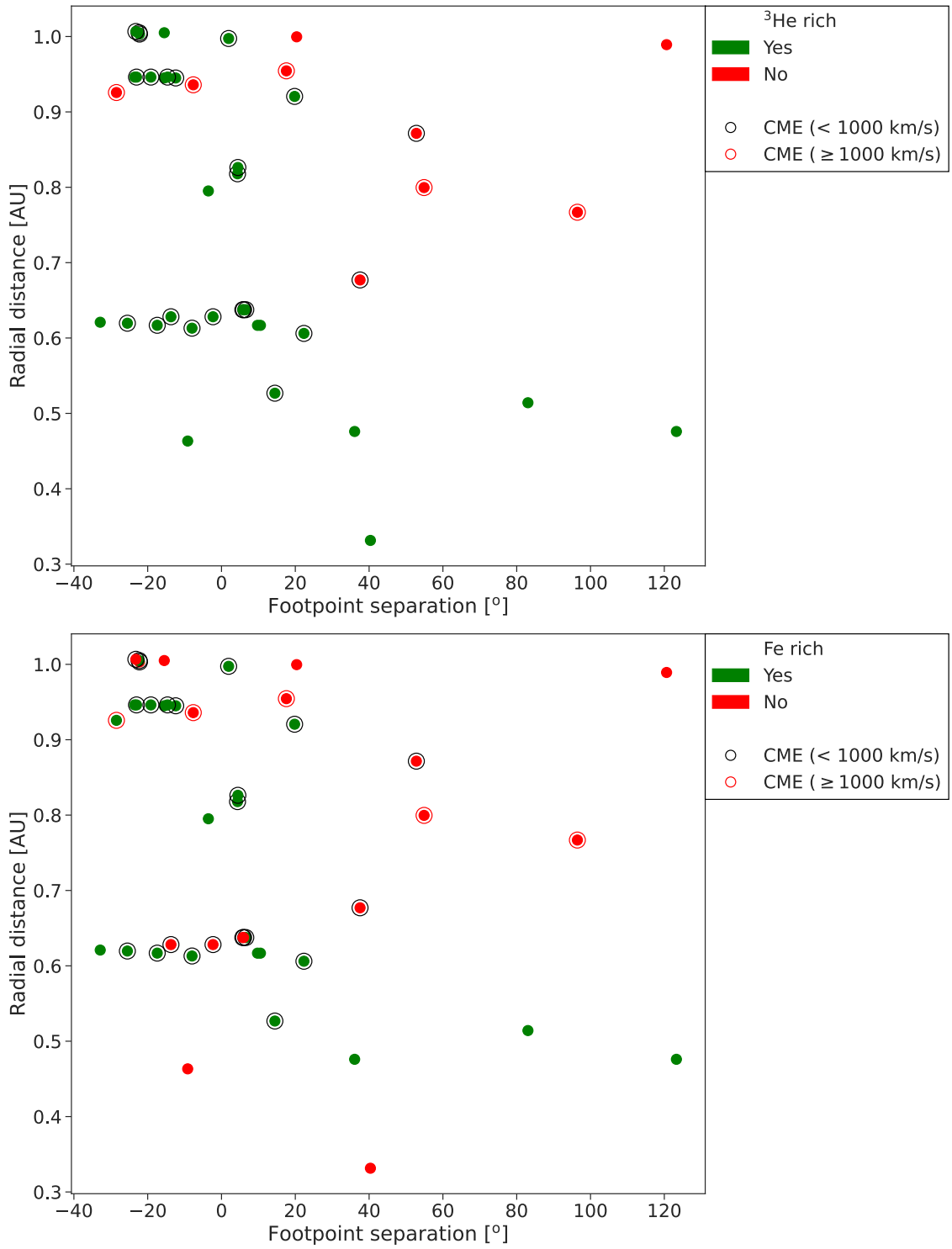


Figure 19: Radial distance from the Sun vs. magnetic footpoint separation from the flare site of Solar Orbiter. Circled in black are slow CMEs and circled in red are fast CMEs. ^{3}He (top) and Fe (bottom) rich events are colored coded in green.

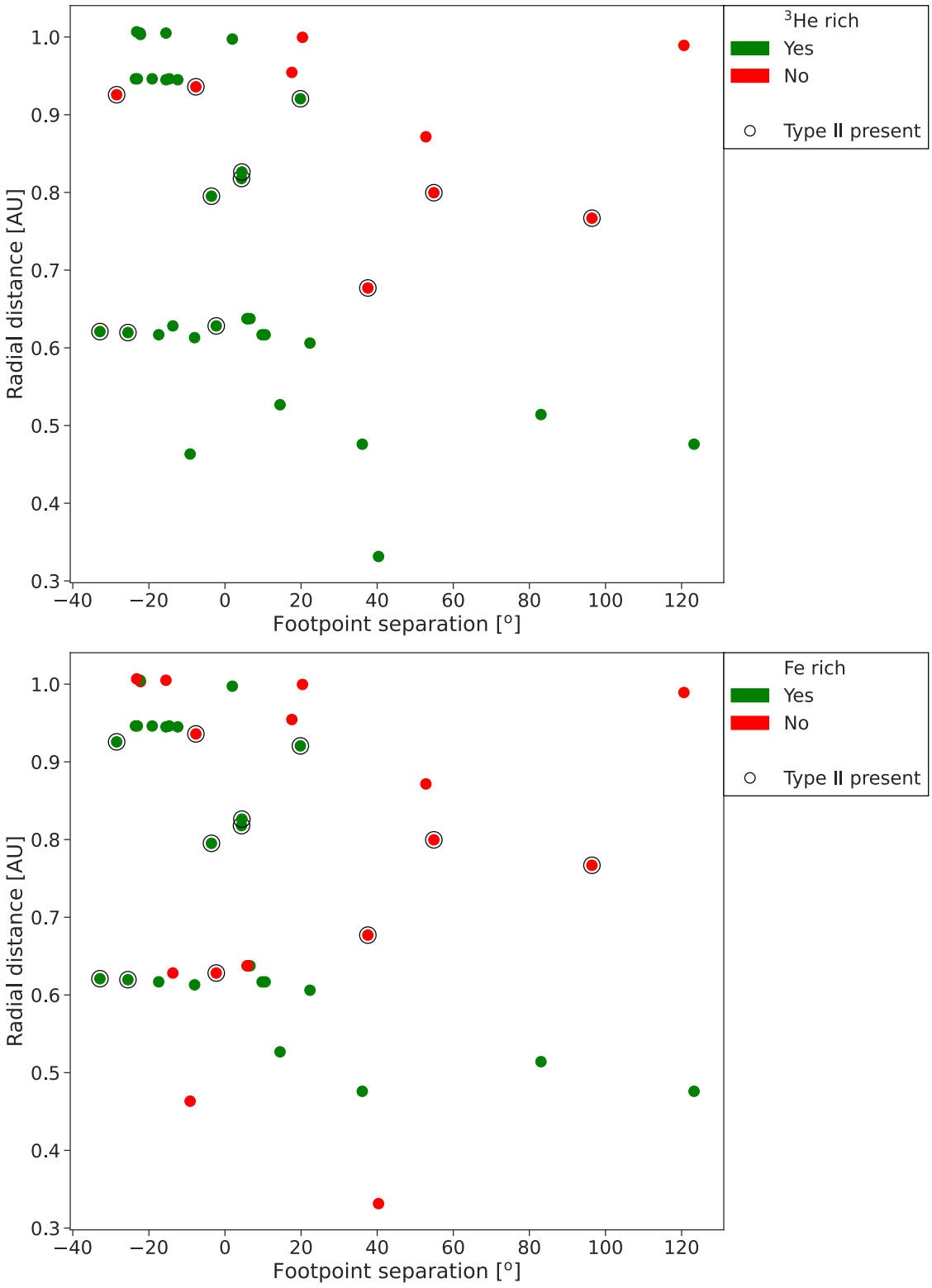


Figure 20: Radial distance from the Sun vs. magnetic footpoint separation from the flare site of Solar Orbiter. ^{3}He (top) and Fe (bottom) rich events are colored coded in green, and black circles indicate the presence of a type II burst.

particle events; the electrons simply don't reach the highest energies without the acceleration capability of a shock. This is also shown in Figures 14a and 14b, where we see most of the events with a type II appearing in the higher energies, as opposed to the events with no type II that often appear to not reach those energies.

The Wind data histogram of Figure 14b appears to have a binomial distribution. While interesting, this is most likely due to the instrument not being as sensitive at around $\sim 13 - 28$ keV*. For the Solar Orbiter histogram of Figure 14a, if a bimodal distribution is forming, it could be because of the switch in instruments from EPT to HET.

The presence of type II bursts, or shock waves, is highly correlated with flare intensity in the Solar Orbiter data. This is to be expected because of the BFS; shock waves accelerate electrons to higher energies and high maximum energies are correlated with strong flares. An interesting observation is that the presence of type IV bursts is even more strongly correlated, especially since the physics of type IVs is not fully understood.

The presence of slow CMEs, fast CMEs, type IIs, type IIIs and type IVs correlated positively and statistically significantly in the Wind data. Even if the correlation is not very strong, like in the case of type III bursts, this observation is in agreement with the BFS notion that a more intense flare will produce an event with more of everything observable.

None of the events that had type IIs associated with them have a spectral index softer than -4, indicated by the dashed line in Figure 17. This could indicate re-acceleration of flare-originated SEEs in shocks or that events with shocks directly lead to harder spectra. As stated earlier, the difference in the two plots appears to be the points that are below this line, that is, those events that have very steep spectra at 43 keV. This set of points is no longer present at 200 keV. This could

*Linghua Wang, private communication

suggest a subset (lower intensity part of the 43 keV plot) of smaller, low intensity events, perhaps those where only a flare contributed to the event or where only the flare contribution of the event is observed. This subset of events may or may not also be seen in the low intensity, low energy parts of the scatter plots of Figure 13.

One would not expect to see the elemental abundance indicators of flare acceleration if an event is only due to the process of shock acceleration. The finding that ^{3}He rich events are anti-correlated with flare intensity plays into this notion, however, Fe rich events not being statistically significantly correlated does not. We also notice that the most separated event in Figures 19 and 20 is both ^{3}He and Fe rich and yet has no CME or shock. The idea that an event could be said to consist only of the flare component if it is ^{3}He or Fe rich is questionable according to the data presented here – neither seem to be a reliable indicator on the basis of which one could conclude that an event consisted only of a flare component. This observation also hints in the direction of re-acceleration in shocks, instead of pure shock events.

There are a handful of uncertainties associated with the work presented here. First is the quite small sample size of the Solar Orbiter data set. This was in part out of necessity to fulfil the selection criteria for the data.

At the time of writing, the Solar Orbiter's STEP and EPT instruments, which measure in part overlapping energy channels, have a mismatch in their measured particle intensities. The energy spectra are the only observables in this thesis that can be affected by this. This problem appears to be, at least in part, due to the instruments having different configurations in their fields of view; EPT sees in a 30° cone-shaped beam, while STEP has a box-like 28° x 54° field of view. The problem is not well documented and cause is not entirely understood, but can be alleviated by omitting data from the edge of STEP's field of view and using only the centermost pixels of the instrument, and by applying a shift to the STEP data to bring it in-line with EPT measurements. Neither of these corrections are applied in the data used

in this thesis because the corrections were not available for all of the events, and so self-consistency within the data was preferred [32, 46].

CME data obtained from the SOHO/LASCO catalogue is possibly subject to projection effects due to its position at the L1 point with no other references from which to calculate CME speeds. This could, in part, explain why the spectral indices in Section 4.1 don't seem to correlate with the CME speeds. The shocks speeds calculated via the program written for this thesis, discussed in Section 3.1, are probably quite accurate even though the process of calculating the speed of the shock from a type II burst is quite dependent not only on the density model used but also on where on the burst one decides to measure the speed from, which can be a subjective process in the absence of radio imaging. Spectra of all of the type IIs detected were not able to be acquired, and for those cases the values from the RTSN NOAA catalogue were used for the shock speeds. Their process is most likely automated, and no guarantees for the correctness of the speeds is given. The shock speeds calculated by hand compared to the speeds from the catalogue did not seem to follow a pattern of being either consistently lower or higher, suggesting that their method simply takes the starting and ending frequency into account, disregarding which lane and/or harmonic number it was a part of.

The spectral index plot at 43 keV (Figure 17, left) suggests that there could be a subset of, perhaps flare only, events with very steep spectra. Same thing can be observed in the 3D plots (Figure 18) where the lowest intensity, lowest energy events that have no type IIs could form a cluster. Attempts to use principal component analysis and clustering algorithms in the data analysis part of this thesis were quite unsuccessful, mainly due to a lack of consistency in the results across different starting parameters and the clustering algorithm used.

6 Conclusions

The primary science question of this thesis was the different acceleration regions and/or processes of solar energetic electrons. This was studied from the viewpoint of flare acceleration, shock acceleration and the possible mixing of the two.

The Solar Orbiter dataset showed the big flare syndrome trend – strong and intense flares usually come with fast and wide CMEs, shocks and radio bursts as well as an energetic SEP event. From the point of view of the science question, this poses the difficulty that in the largest events usually both flare and shock contributions are present, making it difficult to determine the origin of the SEEs. This highlights the importance of case studies of smaller, lower intensity events.

The results shown strongly indicate, despite the BFS, that shocks have an important role in producing the most energetic and intense SEE events. The highest electron energies are reached by the events that have a shock associated with them. This is seen in Figure 13 where the majority of the shocks are seen in the high intensity, high energy events. This is also seen in the histograms of Figures 14a and 14b where the majority of the shock associated events are again in the high energy bins.

The type IV radio bursts also appear to follow the BFS trend in both of the data sets. The trend is not as strong as with type IIs and fast CMEs. It is difficult to draw any parallels to physical phenomena because of the quite elusive nature of the type IVs, so no conclusion is drawn, other than that it is interesting, and worth looking into in the future. The most likely reason for type IVs is trapped electrons inside post-eruption loops behind CMEs. The correlation observed would make sense in this case.

Findings with the spectral indices at 43 keV and 200 keV are in agreement with shock waves in the solar corona accelerating electrons to higher energies and thus to flatter spectra. The observation that none of the shock associated events have a

spectral index softer than -4 at 43 keV demonstrates this. The low intensity, very steep spectrum events at 43 keV that don't reach 200 keV could be considered as a set of events, perhaps those where only a flare component is present. This is further backed up by the lack of type II bursts observed in these events.

Observing increased ^{3}He/^{4}He or Fe/O ratios (at least with the definitions presented here) don't appear to be reliable indicators of the event being observed consisting only of a flare component, meaning that one can't be certain a ^{3}He or Fe rich event does not have an associated shock. The observation that the least magnetically connected event in Figures 19 and 20 is ^{3}He and Fe rich but has no CME or shock illustrates this. The findings hint at re-acceleration of flare particles in shocks, instead of shock-only events.

A natural future prospect is extending the sample size by more events as they get analysed and as the list gets more complete. Future avenues of analysis could include a more detailed look at clustering the data. Also, the observation that the likelihood of seeing type IV radio burst seem to follow more or less the same trend as type II bursts is of interest. Radio imaging and knowledge about the magnetic connection of radio sources could improve the distinction between flare and shock acceleration when type II and III radio bursts are observed.

Acknowledgements

A big thank you goes to Nina Dresing and Diana Morosan for all the help throughout the entire process of making this thesis. A thank you to Annamaria Fedeli for providing the spectral information from her work.

I'd like to thank the entire space research laboratory team for taking me in as one of their own and of the great experience I've had working with them. A thank you to Rami Vainio for being at the helm, and for overseeing this thesis.

Lastly, a big thank you to all of the friends I've made in these last few years of my studies. A special thank you to Christina Humina for the continued support and love.

References

[1] M. Hesse and P. A. Cassak. "Magnetic Reconnection in the Space Sciences: Past, Present, and Future". In: *J. Geophys. Res. Space Phys.* 125.2 (Feb. 2020), e2018JA025935. ISSN: 2169-9380. DOI: 10.1029/2018JA025935.

[2] Donald V. Reames. *Solar Energetic Particles*. Cham, Switzerland: Springer International Publishing, 2021. ISBN: 978-3-030-66402-2. URL: https://link.springer.com/book/10.1007/978-3-030-66402-2.

[3] D. Müller et al. "The Solar Orbiter mission. Science overview". In: *Astron. Astrophys.* 642 (Oct. 2020), A1. ISSN: 0004-6361. DOI: 10.1051/0004-6361/202038467. eprint: arXiv:2009.00861.

[4] NASA. https://www.nasa.gov/image-article/solar-rotation-varies-by-latitude/.

[5] L. G. Bagalá et al. "Magnetic reconnection: a common origin for flares and AR interconnecting arcs". In: *Astron. Astrophys.* 363 (Nov. 2000), pp. 779–788. ISSN: 0004-6361. URL: https://ui.adsabs.harvard.edu/abs/2000A%26A...363..779B/abstract.

[6] M. G. Kivelson and C. T. Russell. *Introduction to Space Physics*. Cambridge, England, UK: Cambridge University Press, Apr. 1995. ISBN: 978-0-52145714-9. URL: https://books.google.fi/books/about/Introduction_to_Space_Physics.html?id=qWHSqXGfsfQC&redir_esc=y.

[7] E. N. Parker. "Dynamics of the Interplanetary Gas and Magnetic Fields." In: *Astrophys. J.* 128 (Nov. 1958), p. 664. ISSN: 0004-637X. DOI: 10.1086/146579.

[8] G. Allen Gary. "Plasma Beta above a Solar Active Region: Rethinking the Paradigm". In: *Sol. Phys.* 203.1 (Oct. 2001), pp. 71–86. ISSN: 1573-093X. DOI: 10.1023/A:1012722021820.

[9] NASA. *NASA Spacecraft Capture an Earth Directed Coronal Mass Ejection*. https://www.flickr.com/photos/gsfc/9558459952.

[10] NASA. *X Class Solar Flare Sends 'Shockwaves' on The Sun*. https://www.flickr.com/photos/gsfc/6819094556/.

[11] Karl-Ludwig Klein and Silvia Dalla. "Acceleration and propagation of Solar Energetic Particles". In: *arXiv* (May 2017). DOI: 10.1007/s11214-017-0382-4. eprint: 1705.07274.

[12] NOAA. *NOAA NWS Space Weather Tool*. https://training.weather.gov/nwstc/spacewx/wmo/tool.html.

[13] Ting Li et al. "Two Types of Confined Solar Flares". In: *Astrophys. J.* 881.2 (Aug. 2019), p. 151. ISSN: 0004-637X. DOI: 10.3847/1538-4357/ab3121.

[14] Yosuke Mizuno. *Plasma Astrophysics Chapter 6: Shocks and Discontinuities*. https://itp.uni-frankfurt.de/~mizuno/plastro/PlasmaAstro-ch6.pdf.

[15] Marcel Goossens. *An Introduction to Plasma Astrophysics and Magnetohydrodynamics*. Dordrecht, The Netherlands: Springer Netherlands, 2003. ISBN: 978-94-007-1076-4. URL: https://link.springer.com/book/10.1007/978-94-007-1076-4.

[16] Gregory D. Fleishman et al. "Solar flare accelerates nearly all electrons in a large coronal volume". In: *Nature* 606 (June 2022), pp. 674–677. ISSN: 1476-4687. DOI: 10.1038/s41586-022-04728-8.

[17] Meiqi Wang et al. "The Solar Origin of an In Situ Type III Radio Burst Event". In: *arXiv* (June 2023). DOI: 10.48550/arXiv.2306.01910. eprint: 2306.01910.

[18] Lewis Ball and D. B. Melrose. "Shock Drift Acceleration of Electrons". In: *Publ. Astron. Soc. Aust.* 18.4 (Jan. 2001), pp. 361–373. ISSN: 1323-3580. DOI: 10.1071/AS01047.

[19] Matthew G. Baring. "Diffusive Shock Acceleration: the Fermi Mechanism". In: *arXiv* (Nov. 1997). DOI: 10.48550/arXiv.astro-ph/9711177. eprint: astro-ph/9711177.

[20] Marc Pulupa. *Space Physics Illustrations*. [Online; accessed 10. Oct. 2024]. URL: https://sprg.ssl.berkeley.edu/~pulupa/illustrations.

[21] George A. Dulk. "Radio Emission from the Sun and Stars". In: *Annual Review of Astronomy and Astrophysics* Volume 23, 1985 (Sept. 1985), pp. 169–224. DOI: 10.1146/annurev.aa.23.090185.001125.

[22] D. B. Melrose. "Coherent emission". In: *Proc. Int. Astron. Union* 4.S257 (Sept. 2008), pp. 305–315. ISSN: 1743-9221. DOI: 10.1017/S1743921309029470.

[23] Christian Monstein. *e-Callisto*. URL: https://e-callisto.org/Data/data.html.

[24] G. Mann et al. "Catalogue of solar type II radio bursts observed from September 1990 to December 1993 and their statistical analysis." In: *Astron. Astrophys., Suppl. Ser.* (1996). URL: https://www.semanticscholar.org/paper/Catalogue-of-solar-type-II-radio-bursts-observed-to-Mann-Klassen/48dbd954f77c81fc6d629af23b26dcf143e4a21f.

[25] D. E. Morosan et al. "Connecting remote and in situ observations of shock-accelerated electrons associated with a coronal mass ejection". In: *arXiv* (Dec. 2023). DOI: 10.48550/arXiv.2312.07166. eprint: 2312.07166.

[26] Diana E. Morosan et al. "Exploring the Circular Polarisation of Low–Frequency Solar Radio Bursts with LOFAR". In: *Sol. Phys.* 297.4 (Apr. 2022), pp. 1–17. ISSN: 1573-093X. DOI: 10.1007/s11207-022-01976-9.

[27] D. E. Morosan et al. "Variable emission mechanism of a Type IV radio burst". In: *Astron. Astrophys.* 623 (Mar. 2019), A63. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201834510.

[28] D. J. McLean and N. R. Labrum. *Solar radiophysics : studies of emission from the sun at metre wavelengths*. Cambridge University Press, 1985.

[29] J. T. Schmelz et al. "Composition of the solar corona, solar wind, and solar energetic particles". In: *Astrophys. J.* 755.1 (Aug. 2012). ISSN: 0004-637X. DOI: 10.1088/0004-637X/755/1/33.

[30] Gennady Minasyants, Tamara Minasyants, and Vladimir Tomozov. "Fe/O ratio behavior as an indicator of solar plasma state at different solar activity manifestations and in periods of their absence". In: *Solar-Terrestrial Physics* 4.1 (Mar. 2018). DOI: 10.12737/stp-41201804.

[31] Rahul Kumar et al. "Preferential Heating and Acceleration of Heavy Ions in Impulsive Solar Flares". In: *Astrophys. J.* 835.2 (Feb. 2017), p. 295. ISSN: 0004-637X. DOI: 10.3847/1538-4357/835/2/295.

[32] J. Rodríguez-Pacheco et al. "The Energetic Particle Detector - Energetic particle instrument suite for the Solar Orbiter mission". In: *Astron. Astrophys.* 642 (Oct. 2020), A7. ISSN: 0004-6361. DOI: 10.1051/0004-6361/201935287.

[33] National Geophysical Data Center. "GOES Space Environment Monitor | NCEI". In: *U.S. Department of Commerce* (Aug. 2009). URL: https://www.ngdc.noaa.gov/stp/satellite/goes/index.html.

[34] Lynn B. Wilson et al. "A Quarter Century of Wind Spacecraft Discoveries". In: *Rev. Geophys.* 59.2 (June 2021), e2020RG000714. ISSN: 8755-1209. DOI: 10.1029/2020RG000714.

[35] R. P. Lin et al. "A Three-Dimensional Plasma and Energetic Particle Investigation for the Wind Spacecraft". In: *Space Science Reviews* 71 (1995), pp. 125–153.

[36] National Oceanic and Atmospheric Administration. *Solar Radio Data*. https://www.ngdc.noaa.gov/stp/solar/solarradio.html.

[37] NASA, The Catholic University of America, Naval Research Laboratory. *SOHO LASCO CME catalog*. https://cdaw.gsfc.nasa.gov/CME_list/.

[38] G. E. Brueckner et al. "The Large Angle Spectroscopic Coronagraph (LASCO)". In: *Sol. Phys.* 162.1 (Dec. 1995), pp. 357–402. ISSN: 1573-093X. DOI: 10.1007/BF00733434.

[39] Jan Gieseler et al. "Solar-MACH: An open-source tool to analyze solar magnetic connection configurations". In: *Front. Astron. Space Sci.* 9 (Feb. 2023), p. 1058810. ISSN: 2296-987X. DOI: 10.3389/fspas.2022.1058810.

[40] Gordon Newkirk Jr. "The Solar Corona in Active Regions and the Thermal Origin of the Slowly Varying Component of Solar Radio Radiation." In: *Astrophys. J.* 133 (May 1961), p. 983. ISSN: 0004-637X. DOI: 10.1086/147104.

[41] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[42] Gareth James et al. *An Introduction to Statistical Learning*. Cham, Switzerland: Springer International Publishing, 2023. ISBN: 978-3-031-38747-0. URL: https://link.springer.com/book/10.1007/978-3-031-38747-0.

[43] Skipper Seabold and Josef Perktold. "statsmodels: Econometric and statistical modeling with python". In: *9th Python in Science Conference*. 2010.

[44] J. Bruin. *FAQ What is complete or quasi-complete separation in logistic regression and what are some strategies to deal with the issue?* URL: https://stats.oarc.ucla.edu/other/mult-pkg/faq/general/faqwhat-is-complete-or-quasi-complete-separation-in-logistic-regression-and-what-are-some-strategies-to-deal-with-the-issue/.

[45] S. W. Kahler. "The role of the big flare syndrome in correlations of solar energetic proton fluxes and associated microwave burst parameters". In: *J. Geophys. Res. Space Phys.* 87.A5 (May 1982), pp. 3439–3448. ISSN: 0148-0227. DOI: 10.1029/JA087iA05p03439.

[46] Annamaria Fedeli. Personal communication.

Appendix A Symbols, units and descriptions of all quantities used in this thesis

Table VIII: Quantities used, their symbols, units and descriptions.

Name	Symbol	Unit	Description
Peak intensity	I_{p}	$[(\mathrm{cm}^2\mathrm{srMeVs})^{-1}]$	Peak electron intensity measured at 43 keV, unless otherwise specified.
Maximum energy	E_{max}	[MeV]	Maximum energy event is still seen at, i.e. is above the background.
Flare intensity	I_{F}	$[\mathrm{Wm}^{-2}]$	Flare intensity measured by GOES.
Spectral index	γ_{i}	[-]	Spectral index at energy i [keV]. Energies of 43 and 200 keV were used.
CME speed	v_{CME}	$[\mathrm{kms}^{-1}]$	Speed of CME in the LASCO catalogue.
Shock speed	v_{shock}	$[\mathrm{kms}^{-1}]$	Speed of shock wave. Inferred from type II burst if able, otherwise from NOAA.
CME present		[-]	Binary variable representing presence of CME.
Type II present		[-]	Binary variable representing presence of type II burst.
Type III present		[-]	Binary variable representing presence of type III burst.
Type IV present		[-]	Binary variable representing presence of type IV burst.
^{3}He rich		[-]	Binary variable representing presence of ^{3}He, i.e. elevated ^{3}He/^{4}He ratio.
Fe rich		[-]	Binary variable representing Fe/O ratio. Fe/O $\sim$ 1 is interpreted as Fe rich.

Appendix B Solar Orbiter events and their parameters used in this thesis

Table IX: Solar Orbiter dataset

Event ID	log flare intesity	log electron intensity	log max E	γ at 43 keV	γ at 200 keV	CME speed	Shock speed	CME present	Type II present	Type III present	Type IV present	^{3}He rich	Fe rich
2105091412	-5.40e+00	5.29e+00	2.01e-01	-7.06e-02	-3.97e+00	6.03e+02	8.09e+02	1	1	1	0	1	1
2105220659	-5.85e+00	5.12e+00	2.01e-01	-2.49e+00	-5.06e+00	4.12e+02	NaN	1	0	1	0	1	1
2105221553	-5.49e+00	4.27e+00	-9.73e-01	-4.42e+00	NaN	NaN	NaN	0	0	1	0	1	1
2105230456	-6.26e+00	4.65e+00	2.01e-01	NaN	NaN	NaN	NaN	0	0	1	0	1	1
2105230940	-5.60e+00	4.89e+00	-6.05e-01	-4.29e+00	-4.29e+00	3.92e+02	NaN	1	0	1	0	1	1
2105231122	-4.92e+00	5.78e+00	2.01e-01	-2.59e+00	-4.64e+00	3.84e+02	NaN	1	0	1	0	1	1
2105231731	-5.64e+00	4.56e+00	NaN	NaN	NaN	5.18e+02	NaN	1	0	1	0	1	1
2107240029	-5.37e+00	4.08e+00	-9.00e-01	-3.92e+00	NaN	5.87e+02	3.45e+02	1	1	1	0	1	1
2108241208	-6.05e+00	3.18e+00	-1.40e+00	NaN	NaN	2.89e+02	NaN	1	0	1	0	1	1
2108241222	-6.05e+00	3.94e+00	-1.01e+00	-3.81e+00	NaN	2.89e+02	NaN	1	0	1	0	1	0
2108241250	-6.05e+00	3.88e+00	-1.18e+00	-6.54e+00	NaN	2.89e+02	NaN	1	0	0	0	1	0
2108261824	-5.51e+00	5.17e+00	2.01e-01	-4.40e+00	-4.40e+00	6.44e+02	NaN	1	0	1	0	1	0
2108262336	-5.40e+00	4.82e+00	2.01e-01	-3.47e+00	-3.47e+00	1.97e+02	4.62e+02	1	1	1	0	1	0
2108280525	-5.15e+00	4.53e+00	2.01e-01	-1.95e+00	-3.01e+00	1.91e+02	1.70e+03	1	1	1	1	1	1
2109241623	-6.40e+00	3.80e+00	-1.21e+00	-5.35e+00	NaN	3.00e+02	NaN	1	0	1	0	1	1
2109261200	-6.62e+00	4.17e+00	-9.00e-01	-5.67e+00	NaN	2.62e+02	NaN	1	0	1	0	1	1
2109271157	-6.35e+00	4.23e+00	-1.01e+00	-2.54e+00	NaN	3.81e+02	NaN	1	0	1	0	1	1
2109280658	-5.80e+00	3.91e+00	2.01e-01	-3.14e+00	-2.64e+00	NaN	5.02e+02	0	1	0	0	1	1
2110090642	-4.80e+00	5.67e+00	2.01e-01	-1.73e+00	-3.84e+00	7.12e+02	7.50e+02	1	1	1	1	0	0
2110281553	-4.00e+00	5.39e+00	1.03e+00	-2.09e+00	-1.87e+00	1.52e+03	1.38e+03	1	1	1	1	0	0
2111010154	-4.80e+00	5.38e+00	1.03e+00	-2.68e+00	-2.68e+00	7.53e+02	7.73e+02	1	1	1	1	1	1
2112040754	-5.62e+00	4.16e+00	-1.01e+00	-5.50e+00	NaN	3.39e+02	NaN	1	0	1	0	1	0
2112050754	-4.85e+00	3.81e+00	-1.64e-01	-2.47e+00	-2.47e+00	6.28e+02	NaN	1	0	1	0	1	1
2112051953	-5.80e+00	3.66e+00	-1.64e-01	-2.27e+00	-2.27e+00	NaN	NaN	0	0	0	0	1	0
2112060602	-6.04e+00	3.99e+00	-1.64e-01	-3.07e+00	-3.07e+00	9.92e+02	NaN	1	0	1	0	1	0
2112311026	-6.34e+00	3.77e+00	-9.37e-01	NaN	NaN	NaN	NaN	0	0	1	0	0	0
2201011404	-6.28e+00	3.46e+00	-9.73e-01	-3.87e+00	NaN	1.67e+02	NaN	1	0	1	0	1	1
2201141342	-5.36e+00	4.31e+00	-8.26e-01	-4.42e+00	NaN	1.34e+03	NaN	1	0	0	0	0	0
2201181755	-4.82e+00	5.14e+00	2.01e-01	-3.79e+00	-3.79e+00	1.01e+03	1.43e+02	1	1	0	1	0	0
2201200639	-4.26e+00	5.19e+00	1.03e+00	NaN	NaN	1.43e+03	3.29e+02	1	1	1	1	0	1
2201292336	-4.96e+00	3.29e+00	-1.08e+00	-3.75e+00	NaN	5.30e+02	NaN	1	0	0	0	0	0
2202082159	-5.28e+00	3.47e+00	-1.21e+00	NaN	NaN	NaN	8.05e+02	0	1	1	0	1	1
2203052357	-5.85e+00	5.21e+00	-8.26e-01	NaN	NaN	2.87e+02	NaN	1	0	1	0	1	1
2203060804	-6.21e+00	3.82e+00	-1.08e+00	-6.41e+00	NaN	NaN	NaN	0	0	1	0	1	1
2203091432	-6.28e+00	3.50e+00	-9.00e-01	-2.21e+00	NaN	NaN	NaN	0	0	1	0	1	1
2203091559	-6.35e+00	4.18e+00	-8.63e-01	NaN	NaN	NaN	NaN	0	0	1	0	1	1
2203102039	-5.55e+00	4.82e+00	-1.64e-01	-3.42e+00	-2.98e+00	NaN	NaN	0	0	1	0	1	0
2203302115	-5.57e+00	4.70e+00	-7.53e-01	-2.11e+00	NaN	NaN	NaN	0	0	1	0	1	0
2207232203	-5.96e+00	4.89e+00	NaN	NaN	NaN	NaN	NaN	0	0	1	0	0	0
2208301833	-4.00e+00	4.45e+00	NaN	NaN	NaN	1.25e+03	3.32e+02	1	1	0	0	0	0
2211110143	-5.35e+00	4.52e+00	NaN	NaN	NaN	NaN	NaN	0	0	0	0	1	1
2211110155	-5.34e+00	5.74e+00	NaN	NaN	NaN	NaN	NaN	0	0	1	0	1	1

Appendix C Scatter plot of the Wind dataset used in this thesis

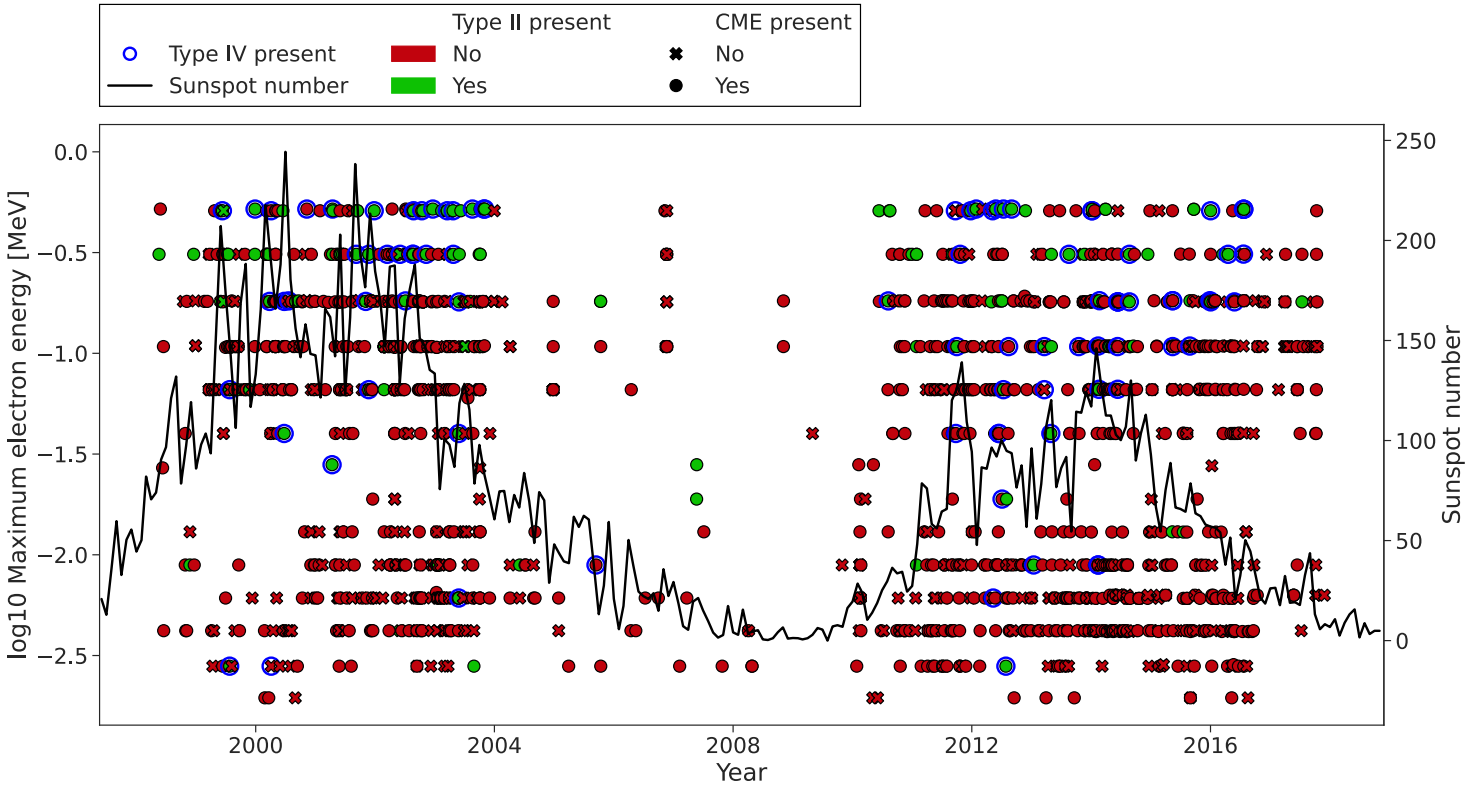


Figure 21: Wind data scatter plot showing all Wind events used in this thesis. Presence of type IIs is color coded: red means no type II, green means one was observed. Crosses indicate the absence of CMEs and circles the presence of one or more. Presence of type IVs is indicated by a circular blue outline. The sunspot number is overplotted in black.