Astronomy & Astrophysics manuscript no. main
June 4, 2017

The Field Distribution of Magnetograms from Simulations of Active Region Formation

S. Dacie[previously at 1], L. van Driel-Gesztelyi[1,2,3], P. Démoulin[2], M. G. Linton[4], J. E. Leake[previously at 4, now at 5], D. MacTaggart[6], and M. C. M. Cheung[7]

[1] Mullard Space Science Laboratory, University College London, Holmbury St. Mary, Surrey, RH5 6NT, UK
e-mail: sallydacie@gmail.com
[2] Observatoire de Paris, LESIA, UMR 8109 (CNRS), F-92195 Meudon-Principal Cedex, France
[3] Konkoly Observatory of the Hungarian Academy of Sciences, Budapest, Hungary
[4] US Naval Research Laboratory, 4555 Overlook Avenue, SW Washington, DC 20375, USA
[5] NASA Goddard Space Flight Center, Greenbelt MD, USA
[6] School of Mathematics and Statistics, University of Glasgow, Glasgow G12 8QW, UK
[7] Lockheed Martin Solar and Astrophysics Laboratory, 3251 Hanover Street Bldg. 252, Palo Alto, CA 94304, USA

Received September 15, 1996; accepted March 16, 1997

ABSTRACT

Context. **The evolution of the photospheric magnetic field distributions (probability densities) was derived for a set of active regions in Dacie et al. (2016). Photospheric field distributions are a consequence of physical processes that are difficult to pin-point only from observations.**
Aims. **In this study,** we analyse simulated magnetograms from numerical simulations, which model the emergence and the decay of active regions. These simulations have different experimental set-ups and include different physical processes, **allowing us to investigate the relative importance of convection, magnetic buoyancy, magnetic twist and braiding for flux emergence.**
Methods. We look specifically at the photospheric field distributions (probability densities **found using kernel density estimation analysis**) and compare the results with those found from observations.
Results. Simulations including convection most accurately reproduce the observed **evolution of the photospheric field distributions during active region evolution.**
Conclusions. This indicates that convection **may play** an important role during the decay phase and also during the formation of active regions, particularly for low flux density values.

1. Introduction

Magnetic flux emergence is an important topic in solar physics, both for its fundamental role in the solar cycle and for its role in eruptive events. Consequently there are strong efforts within the community to improve our understanding of this topic, both from observational and modelling standpoints. Models that are able to reproduce observations can be particularly informative.

There are many different simulations of flux emergence and these include or omit different processes. It is currently not clear which processes are most important in recreating different aspects of flux emergence and active region formation. The most commonly used models of flux emergence use a plane-parallel stratification of the background plasma (*e.g.*, Fan 2001; Murray et al. 2006; MacTaggart & Hood 2009) and insert a magnetic field structure, normally a twisted flux tube, into the simulated convection zone. Models of active region formation that include convection (Cheung et al. 2010; Rempel & Cheung 2014) are quite new and have only a fairly shallow convection zone. Models with a deeper convection zone (Stein et al. 2011, 2012), while able to emulate flux emergence, have not yet managed to self-consistently reproduce spot formation. The recent braid model of Prior & MacTaggart (2016) does not include convection but was inspired by the convective model of Stein et al. (2011, 2012) and inserts a braided field structure, such as those formed by convection in simulations.

The different models focus on different aspects of the emergence, with some aiming to reproduce the small scale structures observed on the photosphere, while others are more concerned with the large scale structures formed in the corona due to the flux emergence. There are also many different characteristics that could be used to judge how well the models represent observations of flux emergence. For example, comparisons can be made with respect to the evolution of the photospheric magnetic field, in particular its spatial extent and organisation as well as its total flux. The amount of twist present in the emerging structure can also be analysed by using magnetic tongues (Luoni et al. 2011), and more generally one can compare maps of injected magnetic helicity (*e.g.*, Démoulin & Pariat 2009, and references therein). Other important characteristics to consider are the formation of realistic sunspots and penumbra (Chen et al. 2017) as well as the amount of interaction with the background coronal magnetic field (*e.g.*, Török 2008; Archontis & Török 2008).

In the study of Dacie et al. (2016), we studied the **distribution (probability density) of the vertical component of the photospheric magnetic field (or flux density) found in** observations of emerging active regions and how this **distribution** evolves over the lifetime of the regions. The results of this previous study are described in Section 2. Here, we aim to use the same analysis on simulated magnetograms, **analysing the vertical field component and comparing** the results with those

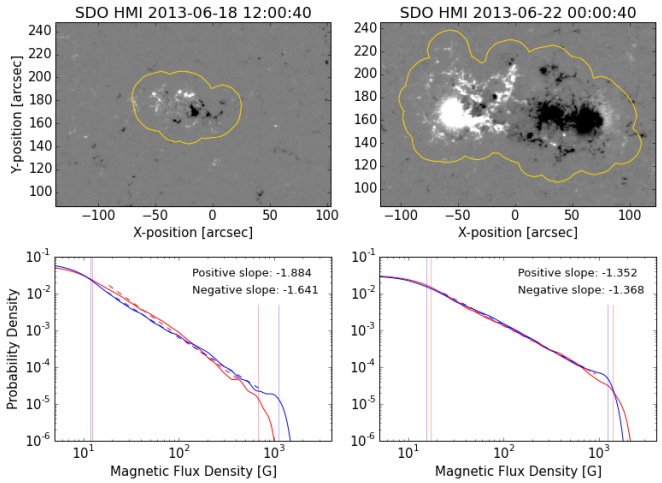

Fig. 1. Magnetograms and their distributions (log-log plot) for NOAA 11776 at the start of its emergence (left) and when spots have formed (right). The distribution of the positive (following) polarity is shown in red and that of the negative (leading) polarity is shown in blue. Dashed lines show the best fit lines and the slope values are given in the legend.

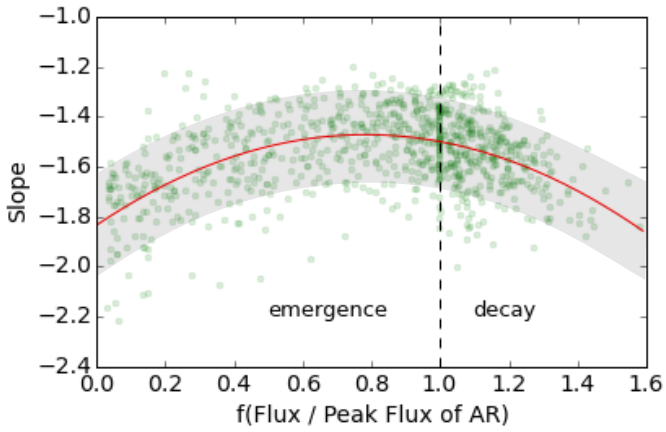

Fig. 2. Observed evolution of the slopes. The emerging and decaying phases are distinguished by the function $f(F/F_{max})$ defined by Equation 1 where F is the magnetic flux averaged between the two magnetic polarities, and F_{max} is its maximum value. The green points show individual distribution slope values obtained during the evolution of the leading and following magnetic polarities for 24 active regions. The red line shows the general trend (second order polynomial least-square fitted to the data points). **The grey shaded area gives an indication of the spread of the data points.** This summarises the main results of Dacie et al. (2016).

of Dacie et al. (2016). Thereby, we can study which processes are involved in active region formation, producing the observed magnetic field distributions. The different simulation set-ups of the analysed models are described in Section 3. **These models allow us to investigate the importance of convection for the emergence and dispersion of magnetic flux, as well as the effects of magnetic twist, braiding and flux tube global curvature on the distribution of the emerging fields.** Alterations made to the analysis method due to the different model set-ups are described in Section 4. We present the results of the analysis in Section 5 and compare and discuss these in Section 6, and summarise the main conclusions in Section 7.

2. Previous Results

The method and results described in this Section are from the observational study of Dacie et al. (2016). We calculated the magnetic field (flux density) distributions of emerging active regions by using kernel density estimation (KDE) analysis (Silverman 1986) **applied to the radialised component of the line-of-sight magnetic field from Helioseismic Magnetic Imager (HMI; Scherrer et al. 2012; Schou et al. 2012) observations. The distributions were plotted on a log-log plot and examples** are shown in Figure 1 for NOAA 11776 near the beginning of its emergence (left) and around the time of maximum flux (right). **The maximum flux was calculated using the radialised field component, and this was done separately for the positive and negative field.** All the distributions were found to have some common features, namely two turning points (indicated by vertical straight lines in Figure 1) and a section between them that could be well approximated by a straight line. The turning points were referred to as the first and second knees and occurred at values of $\sim$ 10 and 1000 gauss respectively.

The slope of the straight line section varied, with values in the range $[-2.2, -1.5]$ at the start of the emergence and rising to $[-1.7, -1.2]$ with the peak slope value on average occurring just before the time of maximum flux. The evolution of the slope is shown in Figure 2, with the evolutionary stage characterised by $f(F/F_{\max})$, where F is the magnetic flux and $F_{\max}$ the maximum flux achieved by the region (as defined in Dacie et al. (2016)). $f(F/F_{\max})$ is designed to separate the emergence and decay phases as

$$f(F/F_{\max}) = F/F_{\max} \quad \text{for } t \le t_{\max} \qquad (1)$$
$$= 2 - F/F_{\max} \quad \text{for } t > t_{\max}\,,$$

where t is the time and t_{max} the time of maximum flux. Figure 2 combines data from 24 active regions, with each point indicating the slope value for a single magnetogram and the general trend shown by the red line. Data from both the leading and following polarities were combined in Figure 2, as no significant difference in slope values was found between the two polarities for the regions studied.

Analysis of a few older regions and quiet Sun regions showed that the slope values continue to decrease after the period shown in Figure 2 towards the quiet Sun value of ~ -3. A simple model of classical diffusion was found to produce a slope of -1 in contradiction to the observations, leading us to conclude that processes other than diffusion, *i.e.* convection, play a key role in active region decay.

In the theory section of our previous study, we also considered the distribution formed by magnetic sources placed below the photosphere, which produce similar distributions and slopes irrespective of their size and number. This suggests that the topology of the photospheric magnetic field does not necessarily affect the slope, so that bipolar active regions may have the same distributions as more complex regions.

3. Numerical Simulations

Seven numerical simulations were analysed. The first of these, published in Cheung et al. (2010); Rempel & Cheung (2014), has a convection zone with a depth of 15.5 Mm and includes convection by solving the MHD and radiative transfer equations self consistently under the assumption of local thermodynamic equilibrium. The simulation is run until a state of statistical equilibrium is reached and then a toroidal flux tube is advected through the base of the computational domain. The data

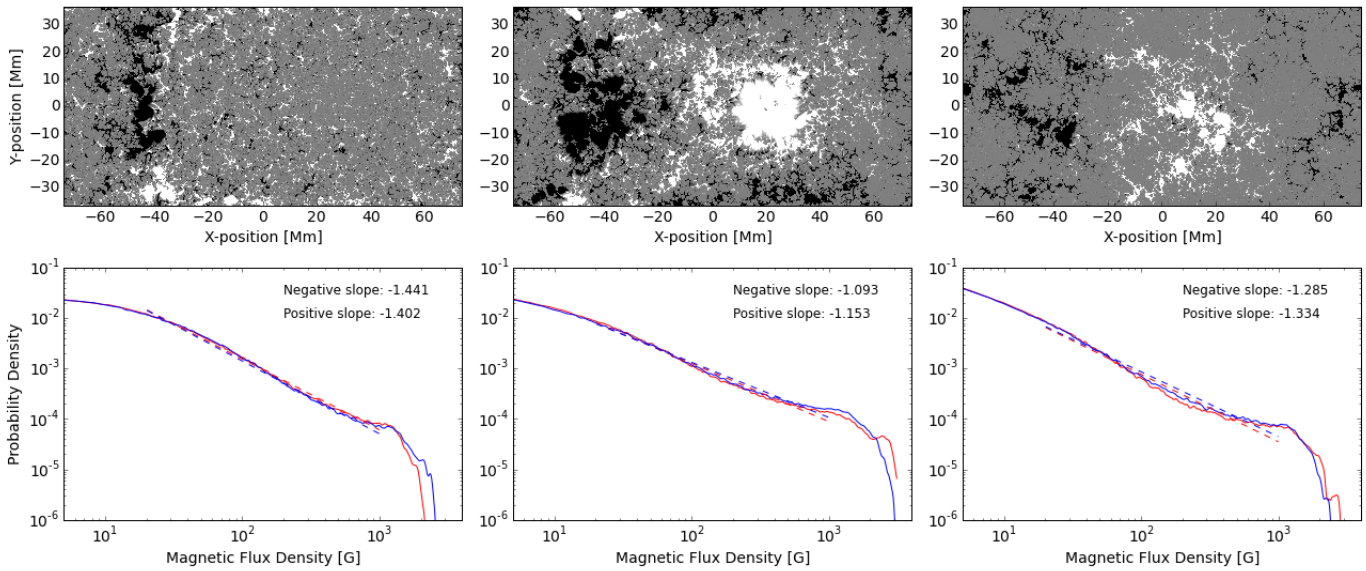

Fig. 3. Simulated magnetograms from Rempel & Cheung (2014) and their distributions. The left hand column shows an example during the emergence phase, the middle at maximum flux and the right during the decay phase at the end of the simulation run. The red line shows the distribution of positive field (the leading polarity) and the blue the negative (following). The best fit lines between 20 and 1000 gauss are shown as dashed lines and their slopes are given in the legend.

we studied comes from a simulation run where the flux tube has no twist and is asymmetric, with a torus-aligned flow to represent the influence of angular momentum conservation as a flux tube rises through the solar convection zone (Rempel & Cheung 2014). **This is the only simulation considered in this study that includes the decay phase, although the simulation run ends when the active region still contains well defined negative and positive polarity regions.**

The next four simulations we analysed use a hydrostatic, stratified background plasma, representing the convection zone, photosphere, chromosphere, transition region and corona. Into this equilibrium, a twisted flux tube (flux rope) is inserted into the convection zone.

Two of these simulations (Leake et al. 2013) use a flux rope with a cylindrical geometry, inserted at a depth of $\sim$ 2 Mm and total radial pressure balance is assumed. The centre of the tube is made buoyant by decreasing the density, while the side boundaries are line-tied, so the ends of the flux tube (unperturbed) remain rooted in the convection zone. The subsequent dynamic evolution produces a rising omega-shaped loop, which emerges through the surface, producing a sheared bipolar surface structure. While this simulation does not include the complete interaction of convection and radiation at the model surface, convective flows are induced beneath the surface **in the wake of the rising flux tube. The induced flows then affect** the buoyancy of the rising tube. Later evolution of this emerging magnetic field produces a sheared coronal arcade and a coronal flux rope above the surface.

The other two simulations with a flux rope have a similar set-up, but they use a twisted flux tube with a toroidal geometry (MacTaggart & Hood 2009; Hood et al. 2012) instead of a cylindrical one, allowing plasma to drain more efficiently, so the flux rope reaches the corona more easily. In addition, the spots reach a maximum separation in this model, which is not the case for that using a cylindrical flux rope.

Finally we analysed two braid model simulations (Prior & MacTaggart 2016), which also use a hydrostatic, stratified background plasma, but include a braided rather than twisted initial magnetic field configuration. The large scale geometry of this magnetic structure is toroidal. Two cases were analysed, one with thick braiding and the other with fine braiding, which had been found to produce very different magnetic configurations in the corona (Prior & MacTaggart 2016).

Although the emerged magnetic field structures produced by the different models have very different morphologies, *i.e.*spatial organisations, this does not necessarily influence the photospheric distributions (as shown for other cases in Dacie et al. 2016). Thus, comparisons between the different distributions should provide additional information compared to analysing the magnetic field spatial organisation during flux emergence.

4. Analysis Methods

We aimed to keep the analysis method as similar as possible to that used in the observational study, however some adjustments had to be made to take account of differences arising from the nature of the simulations. These adjustments are discussed below, but for details of the method as a whole please refer to Section 4 of Dacie et al. (2016).

For the simulation of Rempel & Cheung (2014), the resolution is high and the number of pixels large, **with a horizontal domain size of** 147×74 **Mm^2 and a flux tube with a major radius of** 24 **Mm and a minor radius of** 8 **Mm, which expands to fill the domain during the course of the simulation. As such,** the width of the kernel used to make the KDEs was reduced by a factor of five. **In addition, the large amount of data meant that our statistical analysis could also be performed after reducing the resolution. To reduce the resolution, squares of $n\times n$ pixels are replaced by single pixels with the mean value of the square, and the total number of pixels is reduced by a factor n^2. We investigated three lower resolution cases, with n values of 2, 4 and 8.** The simulation uses periodic boundary conditions on the lateral boundaries, which we thought might af-

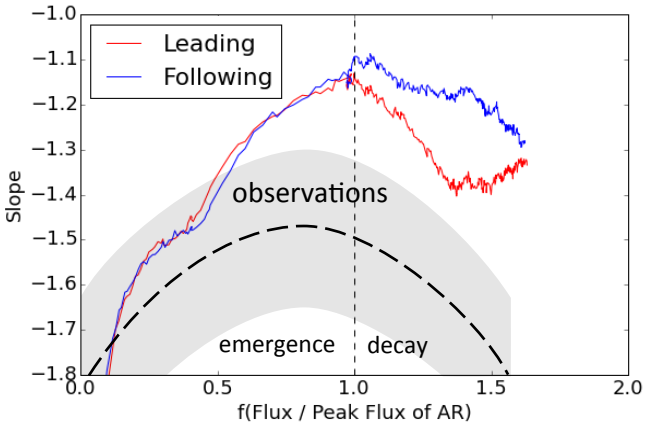

Fig. 4. Evolution of the slope for the simulated magnetograms of Rempel & Cheung (2014). The slope values are plotted against the normalised flux, defined in equation 1, which characterises the evolutionary phase. **For comparison, the observational trend is shown by the thick dashed black line and the approximate range of observed slope values is indicated by the grey shaded region.**

fect the distribution near the edges of the computational box, so we repeated the analysis not including pixels within a distance of 8 Mm (40 pixels) from the edge. Otherwise no area selection procedure was applied to this region, as the active region filled the domain.

For the other sets of simulated data, an area selection procedure was applied. This was very similar to that in the observational study (Section 3.2 of Dacie et al. 2016); the data were first smoothed (using a Gaussian kernel with a width of 7 pixels) and only pixels with smoothed values greater than a certain cut off (20 gauss) were taken. This region was then dilated to include the bordering region within a distance of 8 pixels. If necessary, another dilation followed by an erosion was applied to fill any holes in the selected area. This defines the studied region. Unlike in the observational study where the selection procedure was necessary to remove neighbouring decaying active regions, for the simulated magnetograms this procedure removed a large number of zero (very very small, numerical machine precision) field pixels, which could have influenced the distribution at low B_z values. In summary, we use the original data of the simulation within the defined active region area.

The simulation with a cylindrical flux rope (Leake et al. 2013) uses an irregular grid and this was taken into account when producing the KDEs, with larger pixels making a correspondingly larger contribution to the distribution. The resolution of all the simulated data was higher than data from HMI, however we did not investigate the effects of this for any of the stratified atmosphere simulations, due to the relatively small number of pixels making up the small simulated active region.

5. Results

5.1. Simulation with convection

Examples of simulated magnetograms from Rempel & Cheung (2014) and their distributions are shown in Figure 3. The distributions have a similar shape to those in the observational study, with a roughly straight line section in the middle and a drop off in probability density values at ∼ 1200 gauss. In the observational study, a distinctive turning point (the first knee) was also observed at ∼ 10 gauss (*e.g.*, Figure 1). The distributions from these simulated magnetograms do show a slightly flatter section below ∼ 10 gauss, particularly at early stages of the evolution, however this is not as clear as in the observational study. The best fit straight line was calculated between 20 and 1000 gauss and the evolution of the slope is shown in Figure 4, with the evolutionary stage characterised by $f(F/F_{\max})$ as defined in equation 1.

At the start of emergence, the slopes are steep and negative, they increase to a maximum value of ∼ −1.1 at the time of maximum flux before decreasing again in the decay phase. This behaviour is qualitatively the same as that found in the observational study, but the maximum slope value is slightly greater than those observed, which were in the range [−1.7, −1.2]. Interestingly, the behaviour of the distributions of the two polarities is almost identical in the emergence phase, despite the asymmetry applied to the rising flux tube.

As the resolution of these simulated magnetograms is high and the number of pixels large, we also performed the distribution analysis after reducing the resolution. At half the resolution in both the x and y directions, with pixels of width 384 km, comparable to data from HMI, the slope values showed a small increase (becoming more positive, but generally by less than 0.1). When the resolution was further decreased the slope values decreased, but the difference was still small even for a reduction in linear spatial resolution by a further factor of four. Using a rectangular box to exclude the region nearest the boundary, which was affected by the periodic boundary conditions, also did not have much of an effect on the distributions.

We suggest that the slightly higher (less steep) values of the slopes associated with this simulation compared to the observations (studied in Dacie et al. 2016) could be related to the size of the emerging active region. The simulated region had a peak flux of 1.8×10^{22} Mx, a factor between 3 and 30 larger than the active regions in the observational study. Analysis of a simulation run with a smaller flux tube as well as further observational study of larger active regions and the dependence of the slope values on the peak flux would be needed to confirm this.

5.2. Simulations with a flux rope

Figure 5 shows example magnetograms and their magnetic field distributions from the simulations with a stratified background and a cylindrical flux rope (left) and a toroidal flux rope (right). Both of these simulations are symmetric so the distributions were identical for the positive and negative spots. The distributions found for the two different simulation set-ups are similar, with a roughly straight line section between ∼ 80 and ∼ 700 gauss and a steep drop off after this. The straight line section covers a smaller range than in observed active regions, where it is typically between ∼ 10 and ∼ 1000 gauss. At lower magnetic flux values the distributions have an irregular shape, whose features vary between simulation runs and time steps.

Best fit lines were fitted between bounds k1 and k2, where k1 is defined as the flux density value > 80 gauss with the highest probability density and k2 as the first flux density value moving along the KDE from right to left with a slope value > −2. These values were selected to define the largest approximately straight part of the distribution (in a log-log plot). The evolution of the slopes with time was similar to the observations, starting at steep negative values and increasing over the course of the emergence. However, the first few time steps had unrealistic slope values, not necessarily fitting with this pattern, as the distributions had at best very short straight line sections at these early times. **One possible reason for the discrepancy between these simulations and observations at the first few time steps could**

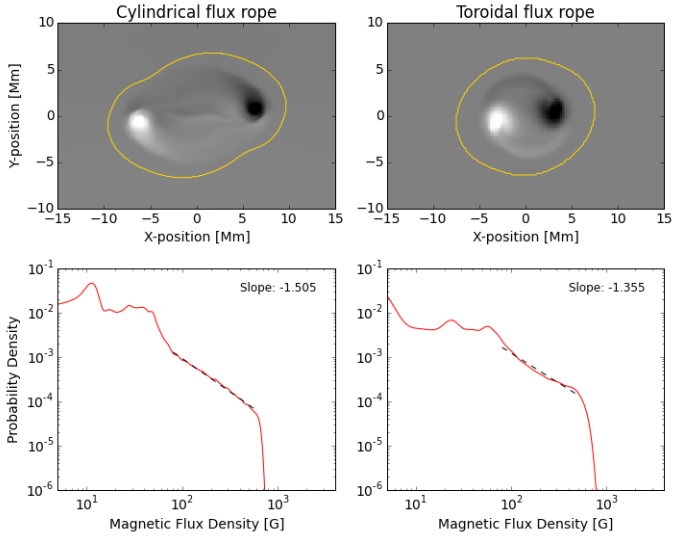

Fig. 5. Magnetograms from simulations using a cylindrical flux rope (left, case C1 Leake et al. 2013) and a toroidal flux rope (right, case T1 MacTaggart & Hood 2009; Hood et al. 2012) and their corresponding distributions. These snapshots were taken near the end of their emergence phases and the magnetograms show a close up view of the emerging regions. The yellow contours outline the area taken for the distribution analysis.

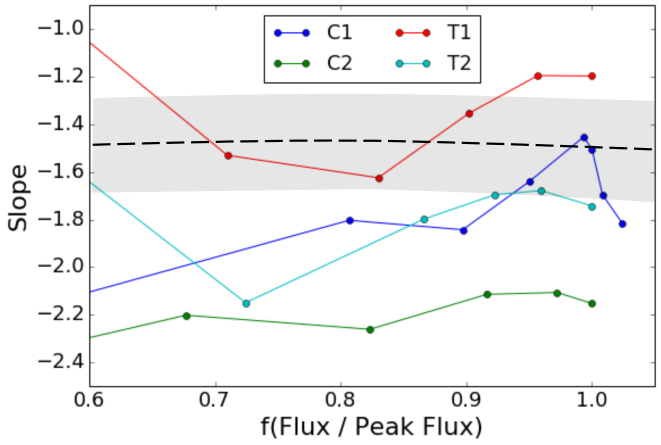

Fig. 6. Evolution of the slope for the magnetograms from the simulations with a flux rope. C1 is the cylindrical case shown in Figure 5 and C2 that with larger flux. **C1 corresponds to simulation SD and C2 to simulation ND of Leake et al. (2013).** T1 is the toroidal case shown in Figure 5, **with initial twist** $q = 0.2/\mathrm{R}$**, and T2 that with stronger twist,** $q = 0.4/\mathrm{R}$**, where** R **is the radius of the flux rope.** The slope values are plotted against the normalised flux, defined in equation 1, which characterises the evolutionary phase. **These simulations do not include the decay phase.** The early emergence is not shown, as the distributions had no clear straight line section at these times. **For comparison, the observational trend is shown by the dashed black line and the approximate range of observed slope values is indicated by the grey shaded region.**

be the requirements for flux emergence in the simulations, which rely purely on magnetic buoyancy for emergence. Another possible explanation is the influence of the strong azimuthal component of the simulated flux ropes, whose role for the vertical field through the photosphere decreases as emergence progresses.

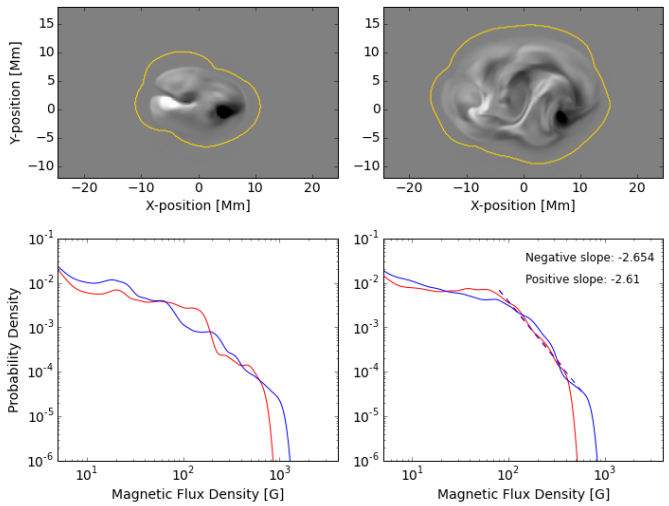

Fig. 7. Simulated magnetograms and their distributions from the thick **(pigtail)** braid model of Prior & MacTaggart (2016). The two time steps are near the beginning (left) and towards the end (right) of the emergence and the magnetograms show a close up view of the emerging region. The red line shows the positive field distribution and the blue line the negative. Slope values are indicated in the legend for the latter time step.

Plotting the slope value against the normalised flux defined in equation 1 (Figure 6), the increase in slope value appeared sudden and close to the maximum flux ($f(F/F_{\max}) = 0.8$-0.9). This could be due to the azimuthal component of the flux rope contributing significantly to the total flux before the region is fully emerged.

Similarly shaped distributions were found from other simulation runs with different parameters. The slopes of these distributions also followed a similar evolutionary trend, but at different slope values. We studied another run using the cylindrical flux rope set-up with a flux rope with a stronger magnetic field (C2 shown in Figure 6) and another run from the toroidal flux rope set-up with a more twisted flux rope (T2). Both the field strength and the twist influenced the slope values, however changes in these parameters did not appear to produce distribution shapes closer to the observed shape (Figure 1) at low flux values.

5.3. Simulations with braided fields

The thick and the thin braid models (referred to respectively as the pigtail and the B4 braids in Prior & MacTaggart 2016) have very different magnetograms. The thick braid produces a swirling pattern of positive and negative flux without strong spots (Figure 7), whereas the thin braid produces two strong polarities (Figure 8) more similar to the flux rope models discussed in Section 5.2. The probability density distributions are also different for the two cases.

Figure 7 shows simulated magnetograms and their distributions at two time steps of the thick braid model. At early times, the distribution appears irregular, with no clear straight line section. A drop off in values occurs at $\sim$ 500 and 900 gauss for the positive and negative polarities respectively. At later stages of the emergence, the shape of the distribution develops **features similar to those in observations**, with a first turning point at $\sim$ 80 gauss, a straight line section up to a few hundred gauss and then a second knee (the drop off). The straight line section, calculated

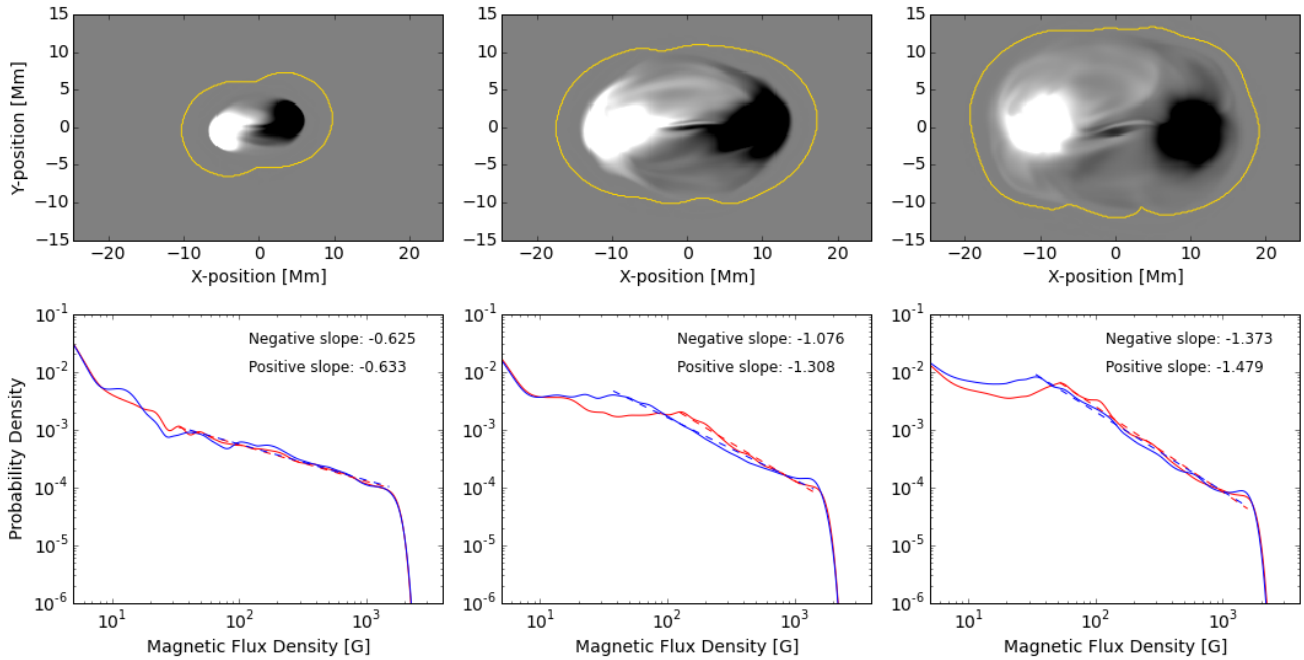

Fig. 8. Simulated magnetograms and their distributions from the thin **(B4)** braid model of Prior & MacTaggart (2016). The three time steps are at the beginning (left), during the middle (middle) and towards the end (right) of the emergence and the magnetograms show a close up view of the emerging region. The red line shows the positive field distribution and the blue the negative. Dashed lines show the best fit lines and their slope values are indicated in the legend.

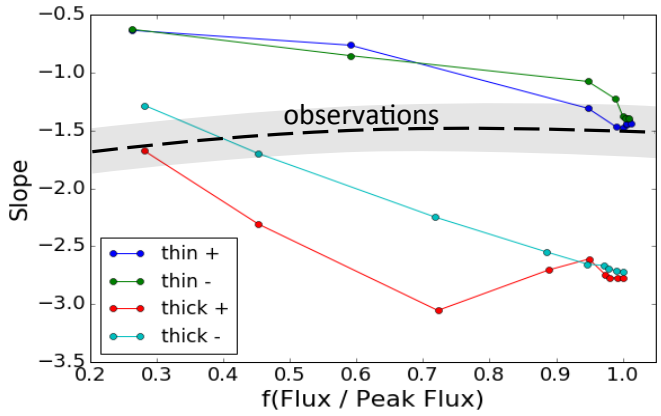

Fig. 9. Evolution of the slope for the magnetograms from the braid model simulations of Prior & MacTaggart (2016). The slope values are plotted against the normalised flux, defined in equation 1, which characterises the evolutionary phase. These simulations are not symmetric and the positive (+) and negative (-) distribution slopes are plotted separately for each of the thick and the thin braid models. **The observational trend is shown by the dashed black line and the approximate range of observed slope values is indicated by the grey shaded region.**

between the bounds defined in Section 5.2, is steep with a slope of ~ -2.6, which may be due to the first knee being at a large B_z value ($\sim$ 80 gauss) and the second knee at a relatively small one ($\sim$ 400 and 600 gauss for the positive and negative polarities respectively). The slopes are steeper than those observed during the emergence phase and the second knee shifting to lower B_z values contradicts observations of active region formation where coalescence of small magnetic loops creates strong spots pushing the second knee to higher B_z values.

Unlike the thick braid model, the thin braid does produce strong spots and results from this model are shown in Figure 8. The distributions show an approximately straight line section and a drop off at $\sim$ 1500 gauss, similar to the observed distributions (Figure 1) except with a higher drop off value. At later times (middle and right columns of Figure 8) a distinct turning point can be seen at the low B_z end of the straight line portion. This is also in agreement with the observations, although this first knee occurs at $\sim$ 40 gauss in this simulation run and $\sim$ 10 in the observed distributions. A best fit straight line was calculated between k1 and k2, with the k2 defined as in Section 5.2 and the k1 as the flux density value $>$ 30 gauss with the highest probability density. The evolution of the slope, shown in Figure 9, was opposite to the observed evolution, with the slope steepening (becoming more negative) during the course of the emergence phase.

6. Discussion

None of the simulations studied here perfectly recreate the magnetic field distributions of active regions on the Sun, but all of them have some similarities to observations. The convective model of Rempel & Cheung (2014) produces distributions with a very similar shape and evolution to the observations. The first knee, at low field strength, is not as clear as in the observations, particularly at later times in the evolution, however the distinctness of the first knee in the observations may be a result of uncertainties related to the magnetic field measurements. The evolution of the slope is very similar to the observations, with a slight difference being that the peak value of the slope for the observed regions occurred just before the maximum flux, slightly earlier than that for this simulated region (compare Figures 2 and 4).

In addition, it would be interesting to let the simulation run for a longer time to see if the slopes from the simulated data

continue to decrease as the active region continues to decay. Towards the end of the simulation run we have studied, the slope for the positive field distribution shows an increase, opposing the expected behaviour, but a longer simulation run may show that this is only temporary. The slope values themselves differ from the observed values, reaching a peak value of ~ -1.1, compared to a value in the range $[-1.7, -1.2]$ for the observed regions (Figure 2). **We suggest that** the difference in slopes may be due to the size of the region, with the simulated region having a maximum flux $\sim$ 10 times that of the **active regions analysed by Dacie et al. (2016), or alternatively** it may be inherently related to the simulation set-up.

Slope values in agreement with the observations were found for the flux rope models **(Figure 6)**, but the distributions produced by these simulations are only a good representation of the observed distributions at middle to large B_z values ($>$ 80 gauss, **Figure 5**). The thick braid model of Prior & MacTaggart (2016) also produced distributions with a straight line section starting at $\sim$ 80 gauss **(Figure 7)**, but in this case the slope was much steeper than observed **(Figure 9)**. Neither the flux rope models nor the thick braid model accurately captured the distribution at low B_z values (up to $\sim$ 80 gauss) or at early times in the emergence.

The thin braid model (Prior & MacTaggart 2016) was found to better reproduce the observed distributions than the thick braid, with a clear straight line section between $\sim$ 40 and 1500 gauss **(Figure 8)**. Compared to the other stratified background simulations, this is closer to the straight line section of the observational data which was found between $\sim$ 10 and 1000 gauss (Dacie et al. 2016), but still not as good as that for the simulated magnetograms from Rempel & Cheung (2014) **(Figure 3)**. The evolution of the slopes for the thin braid model is opposite to the observed evolution **(Figure 9)**. The initially flat slope illustrates the formation of strong polarities much too early in the process, without the coalescence of small flux concentrations.

We expect that the main differences in distribution shape and evolution between the simulations arise as a result of the different simulation set-ups. The differences found in the position of the first knee are not, as one might expect, related to the spatial resolution of the simulations. All the simulations had comparable resolutions, but differing turning point values. Moreover, analysing the Rempel & Cheung (2014) simulations also at lower resolutions did not show much change in the range of the straight line section of the distributions. The first turning point may be related to the necessary field strength and the plasma beta required for the field to break through into the surface. Here the convection simulation (Rempel & Cheung 2014) and the non-convecting simulations (Leake et al. 2013; MacTaggart & Hood 2009; Hood et al. 2012; Prior & MacTaggart 2016) may differ, as the field emerges through a combination of convection and buoyancy in the former, but mainly via the magnetic buoyancy instability in the latter cases. This could explain why the first turning point in the convection simulation occurs at a much lower B_z value ($\sim$ 10 gauss), providing a better representation of the observations for low B_z, than in the non-convecting simulations (where the first turning point occurs at $\sim$ 80 gauss for the flux rope and thick braid models and $\sim$ 40 gauss for the thin braid model).

It would also be important to have an explanation as to why the slopes of the distributions take the values they do. For the simulations with a flux tube (both the convective and stratified atmosphere simulations Rempel & Cheung 2014; Leake et al. 2013; MacTaggart & Hood 2009; Hood et al. 2012), the initial flux tube has a cross section with a Gaussian profile of the axial field strength, which would produce a distribution with a slope of -1 (Section 2.2 of Dacie et al. 2016). Is the initial profile important in producing the photospheric distributions and how is it transformed during the rise and emergence phases?

During the rise phase the flux tube expands, but becomes squashed and concentrated as it arrives just below the photosphere. The effects of these processes could be investigated by analysing the flux tube cross section at different times during its rise for the non-convecting flux tube simulations (Leake et al. 2013; MacTaggart & Hood 2009; Hood et al. 2012).

The field distribution is further transformed during the emergence phase, with horizontal field beneath the photosphere being converted to vertical field crossing it. The magnetic buoyancy instability is responsible for the emergence in the non-convecting simulations, and in the flux rope simulation runs studied here low-wavenumber modes dominate, causing the emergence of one or two magnetic bubbles into the atmosphere. Other non-convecting simulation runs with a lower twist of the initial flux tube show a greater expansion of the tube during the pre-emergence stage, resulting in higher wavenumber modes of the buoyancy instability and a multipolar (more sea-serpent like) emergence (*e.g.*, Hood et al. 2012). The tension associated with the azimuthal field plays a key role in maintaining the flux tube's coherence, affecting which modes are allowed. It is still unclear whether serpentine emergence is due to the effect of vertical flows associated with convection, or due to modes of instability of the sub-surface flux. While previous modelling studies (Hood et al. 2012) have shown that serpentine emergence can be created without convection, granular scale convection increases its presence. We expect that these processes during emergence are important in shaping the distribution, and more so than the initial magnetic field configuration. An indication of this is given by the differing results of the simulations studied here that use similar initial axial field profiles. **Despite this, it is still not clear how boundary and initial conditions affect the photospheric field distributions and additional studies using many more simulation runs would be required to help determine their effects. For example, to investigate whether the initial configuration is important for the distributions,** further analysis could be done using simulation runs with different initial magnetic field configurations of the flux tube.

The twist of the flux tube should influence the distribution, not only through its importance for the mode of the magnetic buoyancy instability and the emergence, but also due to the relative contributions of the azimuthal and axial flux to the magnetogram. A large proportion of the emerged flux comes from the azimuthal component of the twisted flux tube during the early stages of the emergence. Changing the twist of the flux rope in the toroidal flux rope model caused a significant change to the slope values, as did a change in the maximum flux in the cylindrical model (Figure 6). We expect these and other factors to have a strong influence both in these simulations and in others, and further studies could be done to investigate the effects of these parameters. In particular, it would be interesting to analyse a simulation run using the Rempel & Cheung (2014) convective model and a flux tube of lower maximum flux to see if this brings the slope values closer to those observed.

7. Conclusions

Overall, the simulation of Rempel & Cheung (2014) produced the best representation of observations of active region formation in terms of the magnetic field distribution shape, particularly at low flux density values and at the start of the emergence pro-

cesses. This suggests that convective processes play an important role during the emergence phase of active region evolution and especially in areas of relatively low magnetic field strength. **This simulation also mostly reproduced the observed decrease in slope values during the decay phase, leading us to conclude that convection provides a better explanation for active region dispersion than classical diffusion.**

Field distributions in good agreement with observations were also found for the non-convective simulations for flux values > 80 gauss **for the middle to latter parts of the emergence phase.** In this range, buoyancy driven emergence appears to be just as effective at reproducing the observed magnetic field distributions.

Many further studies could be performed to further investigate how the different processes affect the distributions (*e.g.*, studying the distribution at different depths in the convection zone) and how certain parameters, such as twist, influence the distributions. Furthermore, it would be interesting to perform the same analysis on simulations of magnetic flux emergence with a deeper convection zone, which may be considered more realistic.

Acknowledgements. The authors would like to thank Vasilis Archontis for his insightful comments and the referee for their comments and suggestions. SD acknowledges STFC for support via her studentship. LvDG is partially funded under STFC consolidated grant number ST/H00260X/1 and the Hungarian Research grant OTKA K-109276. JEL and MGL were funded by NASA's Heliophysics Supporting Research and Living With a Star Program, and the Chief of Naval Research's 6.1 program. The cylindrical flux rope simulations analysed were performed with a grant of computational time from the DoD's High Performance Computing Program.

References

Archontis, V. & Török, T. 2008, A&A, 492, L35
Chen, F., Rempel, M., & Fan, Y. 2017, Astrophys. J., submitted [1704.05999]
Cheung, M. C. M., Rempel, M., Title, a. M., & Schüssler, M. 2010, Astrophys. J., 720, 233
Dacie, S., Démoulin, P., van Driel-Gesztelyi, L., et al. 2016, A&A, 596, A69
Démoulin, P. & Pariat, E. 2009, Advances in Space Research, 43, 1013
Fan, Y. 2001, Astrophys. J., 554, L111
Hood, A. W., Archontis, V., & MacTaggart, D. 2012, Sol. Phys., 278, 3
Leake, J. E., Linton, M. G., & Török, T. 2013, ApJ, 778, 99
Luoni, M. L., Démoulin, P., Mandrini, C. H., & van Driel-Gesztelyi, L. 2011, Sol. Phys., 270, 45
MacTaggart, D. & Hood, A. W. 2009, Astron. Astrophys., 507, 995
Murray, M. J., Hood, a. W., Moreno-Insertis, F., Galsgaard, K., & Archontis, V. 2006, Astron. Astrophys., 460, 909
Prior, C. & MacTaggart, D. 2016, Geophysical and Astrophysical Fluid Dynamics, 110, 432
Rempel, M. & Cheung, M. C. M. 2014, ApJ, 785, 90
Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, Sol. Phys., 275, 207
Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, Sol. Phys., 275, 229
Silverman, B. W. 1986, Density estimation for statistics and data analysis
Stein, R. F., Lagerfjärd, A., Nordlund, Å., & Georgobiani, D. 2011, Sol. Phys., 268, 271
Stein, R. F., Lagerfjärd, A., Nordlund, Å., & Georgobiani, D. 2012, in Astronomical Society of the Pacific Conference Series, Vol. 455, 4th Hinode Science Meeting: Unsolved Problems and Recent Insights, ed. L. Bellot Rubio, F. Reale, & M. Carlsson, 133
Török, T. 2008, in COSPAR Meeting, Vol. 37, 37th COSPAR Scientific Assembly, 3194