Annual Review of Astronomy and Astrophysics

New View of the Solar Chromosphere

Mats Carlsson,[1,2] Bart De Pontieu,[3,1,2] and Viggo H. Hansteen[1,2]

[1]Rosseland Centre for Solar Physics, University of Oslo, NO-0315 Oslo, Norway; email: mats.carlsson@astro.uio.no

[2]Institute of Theoretical Astrophysics, University of Oslo, NO-0315 Oslo, Norway

[3]Lockheed Martin Solar and Astrophysics Laboratory, Palo Alto, California 94304, USA

Annu. Rev. Astron. Astrophys. 2019. 57:189–226

First published as a Review in Advance on June 21, 2019

The *Annual Review of Astronomy and Astrophysics* is online at astro.annualreviews.org

https://doi.org/10.1146/annurev-astro-081817-052044

Keywords

Sun: atmosphere, Sun: chromosphere

Abstract

The solar chromosphere forms a crucial, yet complex and until recently poorly understood, interface between the solar photosphere and the heliosphere.

- Advances in high-resolution instrumentation, adaptive optics, image reconstruction techniques, and space-based observatories allow unprecedented high-resolution views of the finely structured and highly dynamic chromosphere.
- Dramatic progress in numerical computations allows 3D radiative magnetohydrodynamic forward models to take the place of the previous generation of 1D semiempirical atmosphere models. These new models provide deep insight into complex nonlocal thermodynamic equilibrium chromospheric diagnostics and enable physics-based interpretations of observations.
- This combination of modeling and observations has led to new insights into the role of shock waves, transverse magnetic waves, magnetic reconnection and flux emergence in the chromospheric energy balance, the formation of spicules, the impact of ion-neutral interactions, and the connectivity between chromosphere and transition region.

- During the next few years, the advent of new instrumentation (integral-field-unit spectropolarimetry) and observatories (ALMA, DKIST), coupled with novel inversion codes and expansion of existing numerical models to deal with ever more complex physical processes (including multifluid approaches), is expected to lead to major new insights into the dominant heating processes in the chromosphere and beyond.

Contents

1. INTRODUCTION

This contribution presents a personal view of recent developments in the study of a part of the solar atmosphere called the chromosphere. The term chromosphere comes from the Greek words $\chi\rho\omega\mu\alpha$ (color) and $\sigma\varphi\alpha\iota\rho\alpha$ (ball), alluding to the colored thin rim seen above the lunar limb during a solar eclipse. The color comes mainly from radiation around 656 nm emitted by hydrogen in the transition from the state $n = 3$ to $n = 2$, the Hα spectral line. This fact can be used as a definition: The solar chromosphere is the region of the Sun from where this radiation originates. Without the help of an eclipse, we can block the dominant visible light with a narrow-band filter centered on the Hα line. The appearance of the solar limb through such a filter at different settings is shown in **Figure 1**.

In the wing of the Hα line, we see features protruding outward that are visible as bright features above the visual limb and dark features on the disk. These are called spicules and are covered

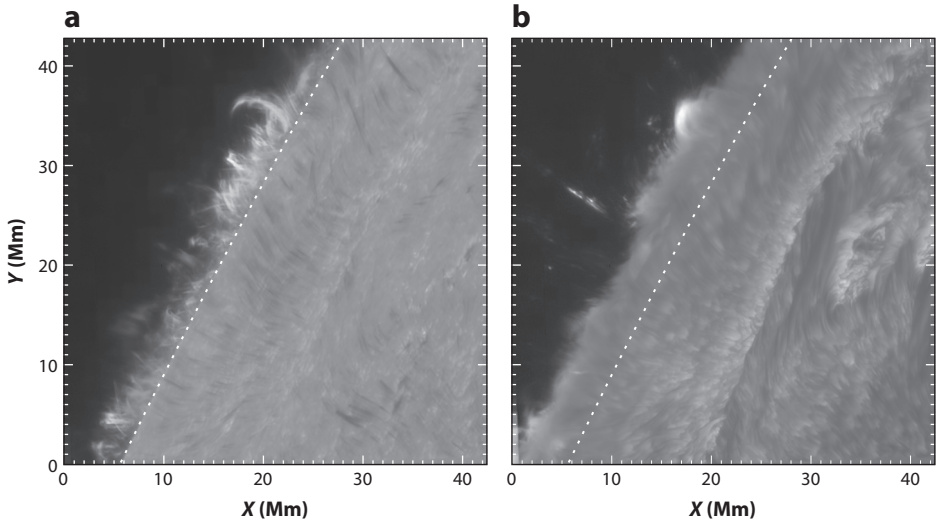

Figure 1

Solar limb as seen through a narrow-band Hα filter. Image in (*a*) the wing of the Hα line at a wavelength corresponding to a blueshift of 41 km s^{-1} and in (*b*) the line core. The visual limb is shown by a dashed line. The observations were made with the CRisp Imaging SpectroPolarimeter instrument (Scharmer et al. 2008) at the Swedish 1-m Solar Telescope (Scharmer et al. 2003) on La Palma.

in Section 5.4. In the core of the Hα line, the opacity is larger and the chromospheric material extends some 5 Mm beyond the visual limb. The opacity in the Hα line is proportional to the number density of the lower level, $n = 2$. At an excitation energy of 10 eV, we need a substantial temperature to get a significant population of this level. If the atmosphere was in radiative equilibrium, we would have a temperature of about 4,000 K in these layers and a population density of the $n = 2$ level of 6×10^{-13} relative to the ground state of hydrogen. In such an atmosphere, we would not see the material above the visual limb in contrast to what is shown in the observations (**Figure 1*b***). It is thus clear that the solar chromosphere is out of radiative equilibrium with a temperature that is higher than that at the top of the photosphere.

Early models were constructed to explain observations in Hα and other strong spectral lines (see Section 3), assuming one-dimensional plane parallel geometry (see Section 4), and they resulted in a temperature drop to a minimum around 4,000 K about 500 km above the visible surface, a temperature rise to 8,000 K at a height of about 2,000 km, and then a very rapid temperature rise to a million-degree corona. These plane parallel models have led to a common notion that the chromosphere is a more or less homogeneous, plane parallel region between these heights. As is abundantly clear in the following, the solar atmosphere above the visible surface is very far from plane-parallel and static; it is inhomogeneous and very dynamic. Instead of talking about a layer, we need a definition based more on physics than geometry: The chromosphere is the region above the photosphere where radiative equilibrium breaks down and hydrogen is predominantly neutral (the latter condition is necessary to distinguish the chromosphere from the corona).

In solar physics, much emphasis has been placed on understanding the physical mechanisms that heat the corona (Reale 2014) and drive the solar wind (Cranmer & Winebarger 2019, in this volume). However, all the nonthermal energy that drives these mechanisms propagates through the chromosphere into the higher regions of the atmosphere. Most of that energy actually remains in the chromosphere, with only a small fraction escaping into the corona or solar wind: While the chromosphere's temperature is only modestly increased over that of the photosphere, its many

Plasma β: the ratio of gas pressure over magnetic pressure

scale heights of dense plasma imply that up to two orders of magnitude more nonthermal energy are required to sustain the chromosphere than the corona. Similarly, all the mass in the corona and solar wind is provided by the dense and partially ionized chromosphere. These chromospheric origins leave intriguing signatures (through, e.g., enrichment of elements with low first ionization potential or FIP) of the complex multifluid physical processes involved in heating plasma to coronal temperatures. Yet many questions remain about how the chromosphere is heated and how its coupling to the corona supplies mass and energy to the outer atmosphere. The role of the partially ionized chromosphere with its intricate ion–neutral interactions in processing the magnetic flux that drives flares and eruptions also remains poorly known. A better understanding of the physical processes in the chromosphere is thus key to making progress on a host of issues in solar physics. Such advances will also be of great interest to astrophysics: The heating of stellar chromospheres remains a major unresolved issue, but perhaps more importantly the chromosphere is a unique laboratory to study many complex physical processes. Because the chromosphere is host to several major transitions, e.g., from fully ionized to partially ionized, from LTE to non-LTE radiative transfer, from strong to weak collisional coupling, and from plasma-dominated to magnetic field–dominated (see Section 2), studies of the chromosphere can provide key insights into universal processes such as wave-mode coupling around the plasma $\beta = 1$ surface, reconnection in a partially ionized atmosphere, nonequilibrium ionization and multifluid interaction effects, etc.

The advances in observations and numerical modeling during the past decade or so have now provided us with a new view of many of these issues. We have recently seen a major improvement in instrumentation capable of exploring chromospheric diagnostics. A wealth of new information on the chromosphere has come from the Swedish 1-m Solar Telescope on La Palma. Larger-aperture observations have come from GREGOR (1.5-m telescope on Tenerife; Schmidt et al. 2012) and the Goode Solar Telescope (previously New Solar Telescope; Goode et al. 2010) of 1.6-m diameter at the Big Bear Solar Observatory. Very soon we will have access to observations from the 4-m-diameter Daniel K. Inouye Solar Telescope (DKIST) on Maui with first light in 2019 and regular scientific operations scheduled to start in 2020. The solar chromosphere can also be observed in millimeter wavelengths, and the Atacama Large Millimeter/submillimeter Array (ALMA; Wootten & Thompson 2009) included solar observations from Cycle 4 in 2016 (with earlier commissioning data). The SUNRISE balloon-borne 1-m solar telescope has provided observations in both broadband imagery and Fabry–Pérot spectropolarimetry during flights in 2009 (Solanki et al. 2010) and 2013 (Solanki et al. 2017). In space, the launch of *Hinode* (Kosugi et al. 2007) in September 2006, and especially its 50-cm-aperture Solar Optical Telescope (SOT; Tsuneta et al. 2008), gave access to seeing-free observations over long time spans. However, in space-based instrumentation, the game changer was the launch in 2013 of the Interface Region Imaging Spectrograph (IRIS; De Pontieu et al. 2014b). It has led to a large expansion of the chromospheric community (and literature) because of the public availability of large amounts of high-quality chromospheric data. Interpreting chromospheric observations is difficult for reasons explained in Section 2, and the development of increasingly realistic computer simulations (see Section 4) in recent years has therefore been equally important for this "New View." It is the combination of advances in observations and numerical modeling that has led to progress.

The topic of the chromosphere is very wide, and it is impossible to cover the whole field, even when giving a personal view rather than an exhaustive review. We thus restrict this contribution to the chromosphere of the quiet Sun and only some aspects of the active Sun, i.e., plage, flux emergence, and spicules. We do not cover prominences, sunspots, or flares.

The structure of this contribution is as follows: In Section 2, we discuss the important physical ingredients under chromospheric conditions; in Section 3, we discuss the diagnostics that provide

information on the chromosphere; in Section 4, we discuss modeling, from semiempirical modeling to comprehensive, "realistic" 3D radiation–MHD modeling; in Section 5, we give our view of the structure, heating, and dynamics of the chromosphere; in Section 6, we discuss connections to the transition region (TR) and corona; and we end with Summary Points and Future Issues.

2. PHYSICS OF THE SOLAR CHROMOSPHERE

To illustrate some of the basic properties of the solar chromosphere, we use the simulation snapshot 385 from the quiet Sun simulation `en024048_hion` of the solar atmosphere made with the code `Bifrost` (Gudiksen et al. 2011). Full data cubes from this simulation are freely available from the Hinode Science Data Centre Europe repository (**http://sdc.uio.no/search/simulations**) and the IRIS modeling page (**http://iris.lmsal.com/modeling.html**), and they are described in detail by Carlsson et al. (2016). Modeling codes, among them the `Bifrost` code, are discussed in Section 4, and the realism of this particular simulation is addressed several times in coming sections. Here, we are interested in basic properties, but we have chosen a 3D model atmosphere for illustration to emphasize the 3D and dynamic nature of the chromosphere.

2.1. Plasma Versus Magnetic Dominance

The dominant factor for the change in physical conditions in the solar atmosphere is the exponential decrease of density with height that comes from hydrostatic equilibrium. The density-scale height is only 100–200 km in the chromosphere, so that even if the density varies by several orders of magnitude at a given height, the exponential decrease with height still dominates (see **Figure 2**). In the photosphere, the magnetic pressure is equal to the gas pressure for a magnetic field strength of about 1.2 kG (0.12 T). This means that outside the strong magnetic concentrations, in strong downdrafts and magnetic structures like pores and sunspots, the gas dominates over the field in the photosphere. Because the magnetic field does not decrease with height as rapidly as the density, the magnetic field dominates in the upper chromosphere. This is illustrated in **Figure 3**. The $\beta = 1$ surface dips down to around $z = 0$ at $x = 8.2$ Mm and at $x = 14.6$ Mm at the locations of magnetic field concentrations in strong downdrafts of cool material. The magnetic field strength at $z = 0$ at these two locations is 0.8 kG and 1.8 kG, respectively. With these exceptions,

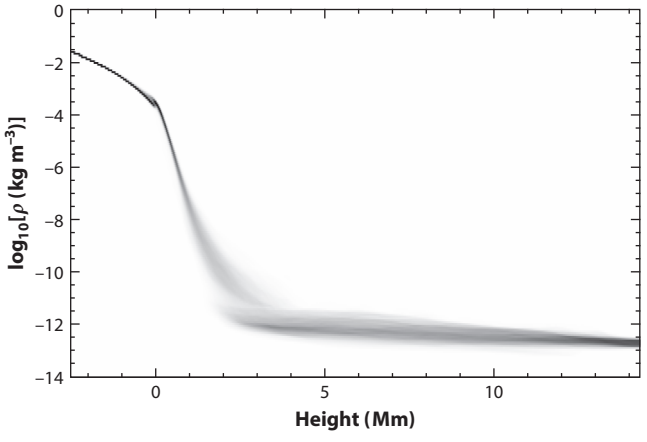

Figure 2

Probability density function of the logarithm of the mass density as function of height in a `Bifrost` simulation encompassing a domain from 2.5 Mm below the photosphere to the low corona.

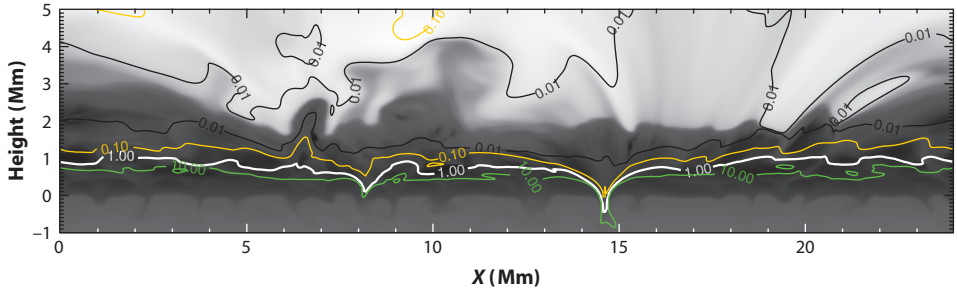

Figure 3

Logarithmic temperature along a vertical slice through a `Bifrost` simulation. The plasma β is shown with contours.

the $\beta = 1$ surface is located in the chromosphere and is quite corrugated, with an average height of 0.9 Mm—this picture is valid for the quiet Sun: In regions of stronger field, such as plage and active regions, the $\beta = 1$ surface is at lower heights. The chromosphere is thus the site of a transition: Plasma motions dominate the magnetic field in the photosphere and low chromosphere, whereas the magnetic field dominates the dynamics in the upper chromosphere. Rapidly falling gas pressure with height implies that the magnetic field also spreads rapidly horizontally, so that already at a few hundred kilometers above the photosphere the field fills essentially all space, even above regions with very weak field in the photosphere.

2.2. Deviations from Local Thermodynamic Equilibrium

The energy transfer in the atmosphere goes from being dominated by convective energy transport in the deep photosphere to energy transport by radiation in the upper photosphere and in the chromosphere. The generation and transport of radiation are set by the emissivity and opacity of the medium, and these depend on the population densities of the atomic energy levels involved in these radiative transitions. The decreasing density with height affects how these population densities are governed. In principle, the population densities are determined by the rate equations

$$\frac{\partial n_i}{\partial t} + \nabla \cdot (n_i \boldsymbol{u}) = \sum_{j,i\neq j}^{n_l} n_j P_{ji} - n_i \sum_{j,i\neq j}^{n_l} P_{ij}, \qquad 1.$$

where n_i is the population density of level i, $\boldsymbol{u}$ is the macroscopic velocity, n_l is the number of levels, and P_{ij} is the transition rate coefficient between levels i and j. The left-hand side represents a continuity equation for the population density of level i, whereas the right-hand side represents a source and a sink term describing the transitions between the levels. When the density is high, the transition rate coefficients are dominated by collisions and are large enough to ensure rapid relaxation to statistical equilibrium (where the left-hand side is zero). We then end up with a linear set of equations in which the solution is given by local thermodynamic equilibrium (LTE), and the population densities are given by Saha's equation for ionization equilibrium and Boltzmann's equation for excitation equilibrium (e.g., Mihalas 1978). This is often a reasonable approximation in the solar photosphere. In the chromosphere, the densities are lower, and radiative rates dominate P_{ij}. We then must solve Equation 1 for all levels together with the radiative transfer equations. Often, statistical equilibrium is assumed, even for chromospheric conditions. The resulting approximation is named non-LTE or NLTE. The literature of NLTE solutions for diagnostic lines is very rich; some examples are given in Section 3.

2.3. Nonequilibrium Ionization

As seen in **Figure 2**, the density is some eight orders of magnitude lower in the upper chromosphere compared with the photosphere. The rates will drop correspondingly (proportional to the density to some power) and will for many processes be slower than typical dynamical timescales. This is the case for the hydrogen ionization balance—a dominant ingredient for the energy and charge balance in the chromosphere. The timescales for hydrogen ionization/recombination in the upper chromosphere can be up to hours, and even in the lower chromosphere timescales are comparable with dynamical timescales (∼10 s; Carlsson & Stein 2002). When waves permeate the atmosphere, timescales are typically short in the compression phase, when the density and temperature are higher, and long in the expansion phase. The consequence is that hydrogen is more ionized than it is in the statistical equilibrium approximation, and temperature amplitudes are larger (Carlsson & Stein 2002, Leenaarts & Wedemeyer-Böhm 2006, Leenaarts et al. 2007).

2.4. Ion–Neutral Effects

Ions and neutrals are often assumed to be strongly collisionally coupled and treated as a single fluid. However, hydrogen is mostly neutral in the solar chromosphere. With decreasing density upward in the chromosphere, collision frequencies are not necessarily high enough to ensure strong coupling between ions and neutrals. A partial coupling between ions and neutrals means that neutrals partly follow the magnetic fields, and ions may partly slip through. To model such a system, in principle one would need to treat each type of particle separately in a multifluid approach.

To simplify the computational approach, and assuming the ions and neutrals dynamics are not completely decoupled, the effects of partial ionization can be dealt with by rewriting the electric field using a generalized Ohm's law that includes ambipolar diffusion (and the Hall term) in a single fluid description (see the sidebar titled The Generalized Ohm's Law). See Ballester et al. (2018) for an overview of the physics of partially ionized plasmas in astrophysical environments and Martínez-Sykora et al. (2015a) for a review of partial ionization effects in the chromosphere.

3. DIAGNOSTICS

We make inferences on the physical conditions in the chromosphere through the analysis of spectral features that form there. It is thus important to assess the available diagnostics.

We can distinguish between two classes of diagnostics depending on the monochromatic optical thickness of the chromosphere. If the chromosphere is optically thin in the specific spectral line, the intensity is set by the background continuum intensity attenuated by the integrated optical thickness and augmented by the integrated emissivity. The width of the line is then set by

THE GENERALIZED OHM'S LAW

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times \left[\mathbf{u} \times \mathbf{B} - \eta \mathbf{J} - \frac{\eta_{\rm Hall}}{|B|} \mathbf{J} \times \mathbf{B} \frac{\eta_{\rm Amb}}{B^2} (\mathbf{J} \times \mathbf{B}) \times \mathbf{B} \right],$$

where $\mathbf{B}$, $\mathbf{J}$, $\mathbf{u}$, η, $\eta_{\rm Hall}$, and $\eta_{\rm Amb}$ are magnetic field, current density, velocity field, Ohmic diffusion, Hall term, and ambipolar diffusion, respectively. The Hall term arises because of decoupling between electrons and ions, whereas the ambipolar diffusion is caused by weak coupling between ions and neutrals. The latter can lead to ambipolar dissipation and heating of plasma.

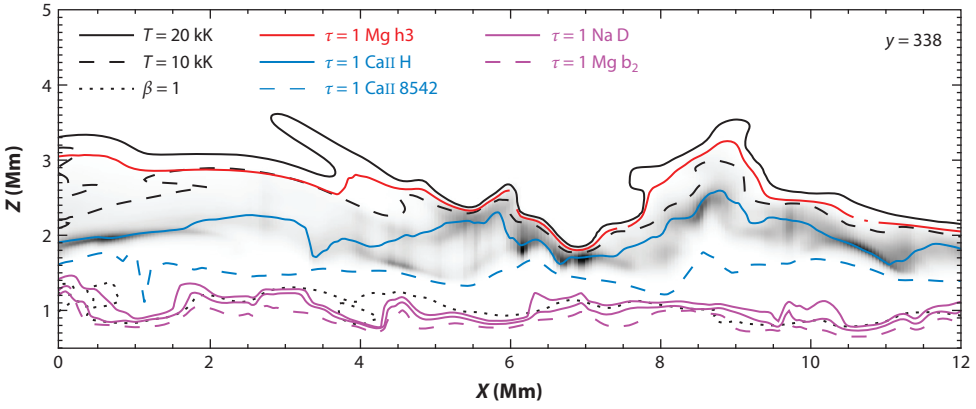

Figure 4

Formation height for various diagnostics in a slice of a `Bifrost` model. For optically thick lines, we give the height of optical depth unity, and for the optically thin line, HeI 1083 nm, the contribution function to the intensity is shown in grayscale. The plasma $\beta = 1$ height is also given.

the thermal width of the opacity profile convolved with the velocity distribution along the line of sight, giving valuable insight into the dynamics of the plasma. The Doppler shift of the line is a measure of the velocity along the line of sight.

If the chromosphere is optically thick in the relevant diagnostics, the intensity is given by the integral of the contribution function, which is often peaked at optical depth unity. In LTE, the source function is equal to the Planck function, and the intensity is then given by the plasma temperature at optical depth unity. Under NLTE conditions, the source function is partly or fully decoupled from the Planck function (i.e., it is normally lower because of photon losses), and the intensity is affected both by the local temperature and by how strongly the source function is coupled to the Planck function.

In the following, we discuss the most common chromospheric diagnostics. The height of formation for various diagnostics is given in **Figure 4**, and a quiet Sun region as seen in several diagnostics is shown in Section 5.

3.1. Continuum Radiation

For the continuum intensity to be formed in the chromosphere, we need an opacity that is high enough to place the optical depth unity in this region. Under solar conditions this happens in the ultraviolet for wavelengths shorter than the bound-free opacity edge of SiI at 152 nm and at radio wavelengths in the free–free continuum from atomic hydrogen and H^- longward of about 0.25 mm. Spectral lines in these wavelength regions are also formed in the chromosphere (or higher). The UV continua were the prime diagnostics used to constrain the VAL (Vernazza–Avrett–Loeser) series of semiempirical models of the solar chromosphere (Vernazza et al. 1981; see Section 4.1). One problem with diagnostics in the UV is that the Planck function depends on the temperature exponentially (in the limit of short wavelengths) such that any average is weighted toward the maximum temperature (Carlsson & Stein 1995). Another problem is that the source function deviates from the Planck function, which makes it necessary to solve a nonlocal, nonlinear, NLTE problem.

The VAL models were also compared with observations in submillimeter and millimeter continua but these observations were at low spatial and temporal resolution. The advantage of

Contribution function: the integrand in

$$I_\nu = \int_0^\infty S_\nu e^{-\tau_\nu} d\tau_\nu = \int_{-\infty}^\infty S_\nu e^{-\tau_\nu} \chi_\nu dz,$$

where I_ν is the intensity at disk center, S_ν the source function, τ_ν the optical depth, χ_ν the monochromatic opacity, and z the height

observations in long wavelengths is that the Planck function varies linearly with temperature and LTE is a good approximation for the free–free processes dominant in the formation of the radiation. The radiation temperature of the intensity is then a direct measure of the plasma temperature. The disadvantage is that the formation height is dependent on the electron density, which to a large extent is set by the nonequilibrium ionization of hydrogen (see Section 2.3). Another disadvantage is the low spatial resolution at a given aperture. The advent of including solar observations for ALMA has changed this situation. The interferometric observations can in principle achieve high spatial and temporal resolution. ALMA may contribute to a range of topics in solar physics (see the review by Wedemeyer et al. 2016). The diagnostic potential has been investigated in detail using synthetic observations from 3D models (e.g., Loukitcheva et al. 2015, 2017a) and observing modes are described by White et al. (2017) and Shimojo et al. (2017a). Early results have been reported (Alissandrakis et al. 2017; Bastian et al. 2017, 2018; Iwai et al. 2017a,b; Loukitcheva et al. 2017b; Shimojo et al. 2017b; Yokoyama et al. 2018), and it is clear that ALMA provides valuable diagnostic information that is complementary to the traditional diagnostics, although several calibration issues have to be overcome to get close to the full diagnostic potential.

3.2. Spectral Lines in the Optical

In the wavelength range of 152 nm to 0.25 mm, we may also find strong spectral lines formed in the chromosphere. In the optical part of the spectrum the most important spectral lines are the CaII H and K lines and IR triplet, the Hα line, and the HeI 1083-nm line.

3.2.1. CaII H and K lines. The CaII H and K lines at 396.8 nm and 393.4 nm have the advantage that they have as lower level the ground state of CaII. In the chromosphere, this is the dominant ionization stage up to a temperature of about 13 kK with significant amounts of CaII up to 20 kK (Carlsson & Leenaarts 2012). The opacity, and therefore the optical depth unity height, is therefore tightly coupled to the column mass. The source function is also partly coupled to the local Planck function. The observed intensity thus gives information on the local temperature. Another advantage is that the short wavelengths give maximum spatial resolution at a given telescope aperture. The disadvantage is the low signal-to-noise ratio because of the low number of photons in the blue, difficulty in getting cameras with high quantum efficiency, and worse seeing in the blue than in the red. As a consequence, we have not seen narrow-band imaging in Fabry–Pérot systems until the installation of the CHROMIS instrument at the Swedish 1-m Solar Telescope (SST) in 2016. There is also a scarcity of polarimetry in these lines, with the notable exception of Martinez Pillet et al. (1990). Partial redistribution (PRD) effects must be included for the proper modeling of the H and K lines. The full 3D problem was addressed by Bjørgen et al. (2018).

3.2.2. CaII IR triplet lines. The CaII IR triplet lines have metastable levels as lower levels, and the opacity and optical depth unity height are therefore more temperature dependent than is the case for the CaII H and K lines. The 854.2-nm line is still an excellent diagnostic line with good sensitivity to magnetic fields. The other two CaII triplet lines, at 849.8 nm and 866.2 nm, are more troubled with blends. All the triplet lines have an unusual shape with increased absorption on the red side of the line core. This shape is due to six components with slightly different wavelengths from six isotopes of calcium (Leenaarts et al. 2014). This isotopic splitting must be taken into account when using the triplet lines for detailed diagnostics.

3.2.3. Hα. The Hα line is the classical chromospheric diagnostic line. The lower level of the line is at an excitation energy of 10.2 eV, which means that the opacity is very temperature dependent.

Because of the low atomic weight, the thermal broadening is also very large and the line is not as good a diagnostic of nonthermal motions as other chromospheric lines. The width of the line is instead a diagnostic of the temperature (Cauzzi et al. 2009). The line core intensity is a diagnostic of formation height because of the strong scattering and 3D effects (Leenaarts et al. 2012). The sensitivity to 3D effects makes the Hα line less suited for inversions but the Hα line gives excellent diagnostics on structure and on energetic events like Ellerman bombs (EBs; see Section 5.3).

3.2.4. He I 1083.0 nm. Also visible from ground-based telescopes is the IR He I 1083.0-nm line. The lower level is a metastable level at the high excitation energy of 19.8 eV. The dominant channel to populate this level is through recombination after photoionization by the coronal radiation field. This process only happens in the upper chromosphere because absorption of the radiation in the hydrogen and helium continua prevents the coronal radiation from penetrating deeper. The He I 1083.0-nm line is a quite unique diagnostic that is optically thin in the chromosphere, sensitive to magnetic fields through the Zeeman and Hanle effects but clearly visible only in the active Sun. With this restriction, this is one of the best diagnostics for determining chromospheric magnetic fields. It also provides detailed structure information in active regions, as evidenced in recent observations with the Goode Solar Telescope (e.g., Wang et al. 2018).

3.2.5. Other optical lines. Several spectral lines in the optical range give information on the upper photosphere and lower chromosphere. Among them can be mentioned the Na I D lines, K I resonance lines (Quintero Noda et al. 2017), Mg I b lines, He I D3 line at 587.6 nm, and MgI 457.1-nm line (Langangen & Carlsson 2009).

3.3. Spectral Lines in the Ultraviolet

The continuum radiation shortward of the Si I edge at 152 nm is formed in the chromosphere, so all spectral lines in this spectral range are also formed in the chromosphere or higher. A few strong lines longward of 152 nm are also formed in the chromosphere, notably lines from abundant elements. The problem with UV diagnostics is that they are only accessible from outside the Earth's atmosphere, and we therefore concentrate on diagnostics for which there is an available observing platform, the IRIS satellite.

3.3.1. Mg II h and k. The resonance lines Mg II h and k at 280.3 nm and 279.6 nm share many characteristics with the Ca II H and K lines, but they are more opaque owing to the abundance of Mg being 18 times higher than that of Ca (Asplund et al. 2009). The ionization balance is also very similar to that of Ca, with Mg II being the dominant ionization stage up to some 14 kK and still 1% Mg II at 20 kK (Carlsson & Leenaarts 2012).

At most locations, the Mg II h and k lines show emission profiles with a self-reversed core [for an overview of Mg II h and k profiles, see Schmit et al. (2015)]. The normal nomenclature is to call the local minimum outside the emission k_{1V} on the violet (blue) side and k_{1R} on the red side, k_{2V} and k_{2R} for the two emission peaks, and k_3 for the central absorption of the k line and correspondingly for the h line. The diagnostic potential of Mg II h and k is the topic of the first two papers in the series on "the formation of IRIS diagnostics" (Leenaarts et al. 2013a,b). At the line core, the source function is so decoupled from the Planck function that the correlation with the local temperature is very weak, but the Doppler shift of the self-reversal is an excellent diagnostic of the line-of-sight velocity at monochromatic optical depth unity. At upper chromospheric densities, optical depth unity is reached after about 100 km along the line of sight with temperatures below 14 kK. The

intensity of the emission peaks shows a correlation with the local temperature at optical depth unity.

3.3.2. Mg II subordinate lines. The equivalents of the Ca II IR triplet for Mg II are the three Mg II subordinate lines. They are in the same wavelength region as the Mg II h and k lines. Two of them are blended, at 279.8 nm, and one is shortward of the Mg II k line at 279.1 nm. The formation of these lines is treated by Pereira et al. (2015). They are normally in absorption but turn into emission if there is a temperature increase at low heights or a large electron density. They become very strong during solar flares.

3.3.3. C II 133.5 nm. IRIS also observes the two strongest chromospheric lines from C II: at 133.45 nm and a blend of two lines at 133.57 nm. The ionization fraction of C II peaks at a temperature of 10 kK and drops rapidly from about 25 kK to reach 10% at 50 kK (Rathore & Carlsson 2015). Depending on the temperature/density structure, the C II lines can be formed both above and below the Mg II h and k lines. The Doppler shift of the core is an excellent diagnostic of the local velocity at optical depth unity, but the intensity and line width are more difficult to use as diagnostics (Rathore et al. 2015a,b).

3.3.4. O I 135.56 nm. The intersystem O I 135.56-nm line is unique. The main contribution to the intensity comes from radiative recombination cascades into the upper level, followed by spontaneous emission. The line is formed under optically thin conditions and its line width is therefore a good diagnostic of the velocity dispersion along the line of sight within the region contributing to the intensity. The intensity is proportional to the electron density squared. See Lin & Carlsson (2015) for details.

3.3.5. Other UV lines. All spectral lines shortward of 152 nm are formed in the chromosphere or higher. There is a wealth of observational data from Solar Ultraviolet Measurements of Emitted Radiation (SUMER) on *SOlar and Heliospheric Observatory* (SOHO), mostly of the quiet Sun, but not much has been done to investigate the diagnostic potential of these lines. One line that is observable with IRIS close to the O I line mentioned above is the C I line at 135.58 nm (Lin et al. 2017).

4. MODELING

Modeling the solar atmosphere entails setting up a set of equations with a certain set of unknowns. When we use a large set of observations to constrain a large set of unknowns, we often talk about semiempirical modeling. Most commonly an equation of conservation of energy is replaced by treating the temperature as a function of height as a free parameter that is to be determined from observations (together with other free parameters). For classical semiempirical models, this optimization was done with manual adjustments without a formal criterion for the goodness of the fit to observations (Section 4.1). With increased computational capacity, automatic adjustment methods were employed. Although there is no formal inversion of the spectra, these methods are traditionally called inversion methods (Section 4.2). A different class of modeling focuses on minimizing the number of free parameters. A given set of physical approximations is used to produce the resulting atmosphere, with little or no input from specific observations (Section 4.3).

4.1. Semiempirical Modeling

It has long been known that the observations in chromospheric diagnostic lines could not be reproduced by radiative equilibrium models. Because the heating mechanisms were unknown, the energy equation was replaced by treating the temperature as a function of height as a free parameter. Even in 1D, the necessary solution of the NLTE radiative transfer equations was very demanding, so the determination of the temperature structure was done by manual trial and error. For an overview of 1D solar model atmospheres, see Rutten (2002). A series of models for the quiet solar chromosphere were constructed in a seminal series of papers (Vernazza et al. 1973, 1976, 1981). The primary observational constraint was provided by continuum observations in the UV band from Skylab. The model corresponding most closely to the average quiet solar chromosphere, often denoted as VAL3C (Variable Atmosphere Laboratory 3C), is the most cited solar chromospheric model. Later models have improved the fit in the temperature-minimum region (Maltby et al. 1986) and removed the need for a temperature plateau to reproduce the hydrogen Lyman-α line (Fontenla et al. 1990, 1991, 1993).

Such 1D, static semiempirical models have been useful in providing numerical laboratories for exploring chromospheric line formation. They have also been used to calculate the amount of mechanical heating needed to balance the radiative losses from the model. However, there are several caveats: A model reproducing the temporal average intensities may give a very different picture than the average of models reproducing the time-dependent intensities. Carlsson & Stein (1995) demonstrated that a dynamic atmosphere with strong shocks gave the same temporal average UV continuum intensities as a VAL model even though the average temperature structure was close to the radiative equilibrium solution. The solution is also not unique—several temperature stratifications may produce the same average intensities. In the age of high spatial and temporal resolution observations, it is a mistake to continue focusing on averages. Instead of trying to reproduce quantities calculated from an average model (like the radiative losses from VAL3C), theoretical models should be tested by confronting detailed synthetic observables (at high spatial and temporal resolution) from the models with the corresponding observables.

4.2. Inversions

In the previous section, we described classical semiempirical models in which the temperature structure was determined by trial and error adjustments. Another class of semiempirical models result from defining an automatic procedure in which we have a formal definition of a norm that measures the deviation between the synthetic observables calculated from the model and the actual observables. An automatic procedure is used to minimize this norm given a number of constraints. Traditionally this type of semiempirical modeling is called inversion even though we do not perform a formal inversion.

Inversion methods make it possible, in principle, to infer the 3D thermodynamic properties and the 3D magnetic field from observations. There are several issues: the instrumental point spread function (PSF) mixing signal from various locations into any given pixel; scattering in the radiative transfer making a pixel-to-pixel inversion invalid; and more free parameters than observables making the problem ill-conditioned. See de la Cruz Rodríguez & van Noort (2017) for a recent review of such issues and recent developments. The newly developed `STiC` code (STockholm inversion Code; de la Cruz Rodríguez et al. 2016) uses the RH NLTE radiative transfer code (Uitenbroek 2001) together with new ideas for sparse inversions of Stokes profiles (Asensio Ramos & de la Cruz Rodríguez 2015). The methods have been validated using classical semiempirical atmospheres but also using a 3D snapshot from `Bifrost`. Such method development using inversion methodology and 3D forward models, also including artificial intelligence and deep learning, shows great promise for revolutionizing the interpretation of chromospheric observations.

4.3. Forward Modeling

The modeling approach for the chromosphere incorporates lessons learned and builds on the major progress that has been achieved in modeling the photosphere. One example of a very successful set of forward models of the photosphere is convective/radiative equilibrium 1D plane parallel models. The energy equation contains radiative energy transport and a mixing-length description of convection. The equation of state includes all relevant elements and ionization stages with ionization and excitation from the assumption of LTE. Such 1D radiative/convective equilibrium models have only three free parameters: the effective temperature, acceleration of gravity, and elemental abundances. The differences between synthetic observables from such a model and observations give information on missing physics or invalid approximations. Strong lines typically show deeper absorption cores than those predicted by these 1D equilibrium models, which is a difference that comes from deviations from LTE. Too narrow line widths come from the neglect of convective motions and the correlations between temperature and velocity (but this informative difference is often masked by introducing another free parameter, microturbulence).

3D hydrodynamic modeling of solar convection was pioneered by Nordlund (1982), who approximated the radiative energy balance through appropriately defined averages of the source function and the opacity. As few as four bins (in contrast to thousands of frequency points in a brute-force approach) in such a multigroup opacity approach gives a decent approximation to the nongray radiative transfer. Detailed comparisons between such 3D solar models and high-resolution spectra show remarkable agreement without invoking microturbulence or other free parameters (Pereira et al. 2013a). Such models are now a standard tool in spectral line synthesis from cool stars for determining stellar abundances.

Several codes have been developed to model hydrodynamics or magneto-hydrodynamics (MHD) in 3D under solar photospheric conditions, and a comparison between the `CO5BOLD`, `MURaM`, and `Stagger` codes was published by Beeck et al. (2012). The extension of these modeling codes to chromospheric conditions is nontrivial because the physical conditions are different (see Section 2). The code that has come closest to implementing these physical conditions is `Bifrost`, which is a new implementation of the methods previously implemented in the `Oslo Stagger Code` (OSC) (Hansteen et al. 2007). Much effort currently goes into also extending other codes to enable chromospheric modeling.

We use `Bifrost` models extensively in Section 5, and we therefore give some details here, concentrating on the special features implemented to enable chromospheric simulations. For a full description, we refer the reader to Gudiksen et al. (2011). The multigroup opacity method of Nordlund (1982) has been extended to account for scattering [Skartlien 2000; see Hayek et al. (2010) for details]. Chromospheric radiative losses in strong spectral lines are calculated in NLTE using simplified recipes (Carlsson & Leenaarts 2012). Optically thin radiative losses are treated using tables calculated from atomic data in CHIANTI, version 5 (Dere et al. 1997, Landi et al. 2006). Thermal conduction is included using operator splitting with an implicit formulation based on a multigrid method. `Bifrost` is an explicit code with diffusive terms in the equations to ensure stability. The diffusive operator employed is split in a small global diffusive term and a location specific hyper diffusion term [see Gudiksen et al. (2011) for details]. Nonequilibrium ionization of hydrogen may be included following the description by Leenaarts et al. (2007) based on the approximations by Sollum (1999). It is also possible to include nonequilibrium ionization of helium [details are given in Golding et al. (2014, 2016)]. The treatment of nonequilibrium ionization of hydrogen and helium is critical for some problems but increases the computational cost by a factor of 2–5. For an early exploration of parameter space, many models have therefore been run without the nonequilibrium ionization modules. Ion–neutral effects can be included in the generalized

Ohm's law formalism (see Section 2.4). This module has so far been used for 2.5D simulations and without simultaneously treating the nonequilibrium ionization effects.

5. CHROMOSPHERIC STRUCTURE, HEATING, AND DYNAMICS

As explained in Sections 1 and 2, the chromospheric structure is set by the magnetic field in an atmosphere in which the density decreases exponentially with height. This magnetic field is created by the solar dynamo, both a global dynamo in the interior and, possibly, a local dynamo in the convection zone (e.g., Hotta et al. 2015, and references therein), with subsequent flux emergence into the chromosphere. The flux emergence happens both at large scales, creating active regions with sunspots, and all over the solar surface in granular scale flux emergence.

Observations at high spatial and temporal resolution reveal a fascinating array of structures and temporal behavior. Historically, in an attempt to bring order to this bewildering zoo of phenomena, names have been given based on the appearance and observed properties (rather than physical mechanisms). We use some recent high-quality images to illustrate some of this richness and then use numerical simulations to try to interpret what the diagnostics tell us about chromospheric structure, heating, and dynamics.

Figure 1 shows the chromosphere as seen in the Hα line toward the limb. From the line-center image (**Figure 1*b***), chromospheric material clearly extends to 3–5 Mm above the visible limb, which is well beyond the extent in plane parallel models (where the corona starts some 2 Mm above the limb). It is also clear that the chromospheric plasma is very organized. Close to the active region in the right part of the figure, we see short, linear features. These are dynamic fibrils with the abrupt end in the upper part caused by heating to coronal temperatures where hydrogen is fully ionized and, hence, there is no opacity in the Hα line. These dynamic fibrils show parabolic motions caused by shocks propagating upward along the magnetic fields (Hansteen et al. 2006, De Pontieu et al. 2007a). The longer linear features shown in the wing of Hα (**Figure 1*a***) as dark features on the disk and bright features off limb are jets with various names given depending on the properties (see Section 5.4).

Another view of the solar chromosphere is provided in **Figure 5** with images taken with the CHROMIS and CRISP instruments on SST. The field of view covers a quiet Sun region. The magnetic field is concentrated by the convective flows into intergranular lanes (seen as dark lanes in the continuum image in **Figure 5*a***) and into a larger pattern, called the network (**Figure 5*b***). In the internetwork regions, the field is weaker but there is weak magnetic field practically everywhere. From the stronger patches of negative polarity vertical field, we see fibrillar structure in the line core images (Hα, Ca II 854.2 nm, and Ca II K; **Figure 5*d–f***) connecting to positive polarity patches. These fibrils are filling the field of view in Hα and Ca II K. The Ca II 854.2-nm line is less opaque and shows fibrillar structure lower down (**Figure 5*e***). These are not always aligned with the upper fibrillar systems. The fibrils are mostly aligned with the magnetic field, although there is observational evidence that this is not always the case (de la Cruz Rodríguez & Socas-Navarro 2011), which may come about because of decoupling between ions and neutrals (ion–neutral slippage; Martínez-Sykora et al. 2016a). At the wavelength of the K_{2V} feature of the Ca II K line (**Figure 5*c***), one can see many bright blobs in the internetwork regions that are caused by acoustic shocks (Section 5.1).

5.1. Quiet Sun

Acoustic waves and shocks play an important role in the dynamics and energetics of weak-field regions. Spectra taken in the Ca II H and K lines in internetwork regions show a typical sawtooth Doppler-shift pattern of the line core coupled with brightenings of the blue peak repeating with

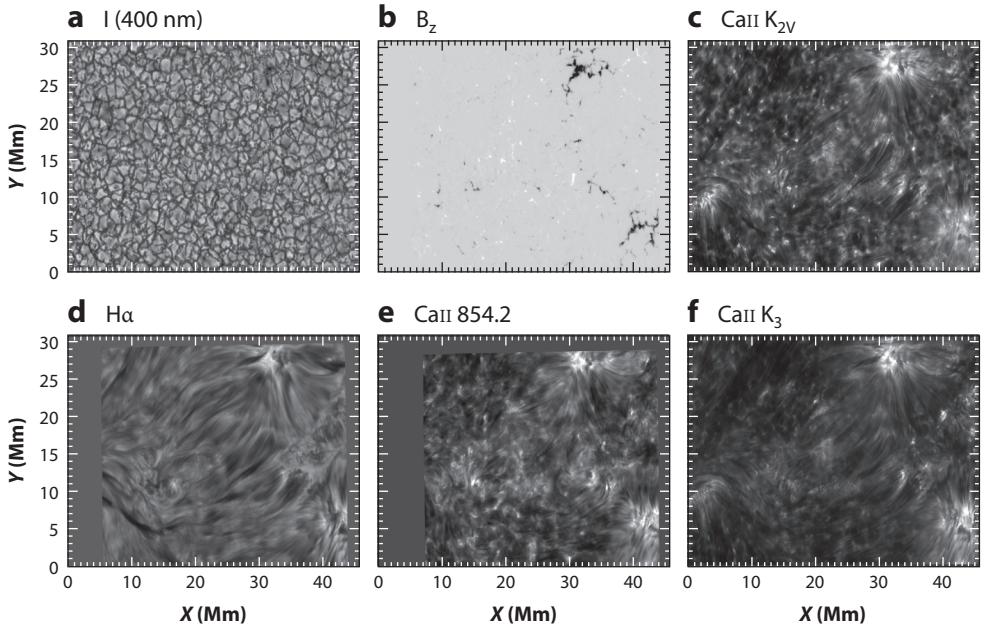

Figure 5

Overview of observations taken by CRISP/CHROMIS on October 12, 2016, at 10:53:12 UT. (*a*) Continuum at 400 nm; (*b*) B_z as inferred from a Milne–Eddington inversion of the FeI 603.1-nm and 630.2-nm lines; (*c*) CaII K image at 13 km s^{-1} to the blue of line center (position of the K_{2V} feature in the mean profile); (*d*) Hα at nominal line center; (*e*) CaII 854.2 nm at nominal line center; (*f*) CaII K at nominal line center.

a periodicity of about 3 min. These spectral features were very successfully reproduced by 1D hydrodynamic simulations of upward propagating acoustic waves shocking in the chromosphere (Carlsson & Stein 1992, 1997).

These simulations of Carlsson & Stein did reproduce the CaII H spectra remarkably well but failed to reproduce the emission lines formed in the upper chromosphere that are observed by the SUMER spectrograph onboard SOHO to have emission all the time. The speculation in Carlsson & Stein (1997) was that these lines are formed above the $\beta = 1$ height where the magnetic field is important, even above photospheric areas with very weak field. Another possible explanation is that these simulations were driven with the observed velocity field in low-cadence observations, thus lacking high-frequency acoustic waves.

The importance of high-frequency waves was tested by Fossum & Carlsson (2005b), who used high-cadence (13 s), seeing-free, intensity observations from the TRACE (*Transition Region and Coronal Explorer*) satellite in the 170-nm and 160-nm bands (formed around 360 km and 430 km height, respectively; Fossum & Carlsson 2005a). NLTE effects and smearing of high-frequency signals because of an extended region contributing to the intensity was handled through calibrations using 1D radiation hydrodynamic simulations. The conclusion was that the observations gave an upper limit to the acoustic flux of 440 W m^{-2}, a factor of ten less than is needed to balance the radiative losses in the semiempirical model VAL3A. More importantly, the 1D simulation with the upper-limit wave spectrum as input did not reproduce the ubiquitous weak emission lines in the UV—the temporal evolution was dominated by the low-frequency acoustic waves already present in the earlier simulations of Carlsson & Stein (1992, 1997).

The TRACE observations have limited spatial resolution (1 arcsec), and it has been speculated that much of the acoustic power is hidden at smaller scales. Observations with *Hinode*/SOT at 0.16-arcsec resolution failed to show a significant increase in acoustic power at small spatial scales (Carlsson et al. 2007).

However, the quantification of acoustic flux in the lower chromosphere is far from settled. Bello González et al. (2010) find significant acoustic flux of 6,400–7,700 W m^{-2} at a height of 250 km using observations from SUNRISE, whereas Beck et al. (2009) find a much lower acoustic energy flux. Beck et al. (2012, 2013a,b) also make inversions of CaII H profiles and find that the quiet Sun does not exhibit a clear chromospheric average temperature rise.

The signatures of acoustic waves can be seen in all chromospheric diagnostics as sawtooth behavior in the CaII lines (because of the temporal variation of velocity and thus Doppler shift associated with shocks) but also as intensity brightenings in UV continua observed by SUMER (Carlsson et al. 1997). These brightenings are also very clearly visible in IRIS slit-jaw images, and Martínez-Sykora et al. (2015b) conclude that the dominant contribution comes from the continuum although there are also sometimes spectral signatures in the CII and SiIV lines, indicating that some fraction of the shock waves may reach all the way up to the TR. Signatures of waves propagating from the chromosphere up to the TR have also been seen in SUMER observations (Wikstøl et al. 2000).

However, there is often also an absence of a clear shock signature in diagnostics formed in the upper chromosphere, leading to the question, "Where have all the shocks gone" (Steffens et al. 1997)? The explanation may lie in the fact that the magnetic field is volume filling in the upper chromosphere, and acoustic waves undergo mode conversion, refraction, and reflection where the Alfvén speed and sound speed are equal (close to the $\beta = 1$ surface) (Bogdan et al. 2003). Depending on the magnetic field topology, the acoustic waves may be completely refracted/reflected or may reach the upper chromosphere and TR.

Although the impact of acoustic waves seems diminished in the upper chromosphere, the spreading of the magnetic field at heights greater than 1 Mm implies that the magnetic field assumes an important role, also in the quiet Sun, not only as an agent for wave conversion but also as a source of heating in its own right. Motions in the convection zone and photosphere stress the magnetic field or cause the emergence of field into the outer solar atmosphere and interactions with the already present ambient field. As gradients in the field structure build up stresses, these will be relieved either violently through reconnection or more gently via Joule heating. Much progress in understanding this "magnetic braiding" (i.e., the "nanoflare" model; Parker 1988) has been made in the past decade, not least due to great improvements in the realistic modeling of this interaction. Such models, in which photospheric motions do work on the magnetic field leading to heating, produce chromospheric heating and coronae of 1 MK or greater temperatures as well as synthetic observables that reproduce many of the properties seen in chromospheric, TR, and coronal diagnostics. However, there are also significant discrepancies between the observations and the simulations as described below.

5.1.1. `Bifrost` models of the chromosphere. The above discussion shows that an understanding of the magnetic field is necessary to model the chromosphere. This also means that one must consider atmospheric models that are fully three dimensional. One such model is the published model `en024048_hion` (Carlsson et al. 2016). This model covers a 24 Mm × 24 Mm area going down 2.4 Mm below the average height of $\tau_{500\mathrm{nm}} = 1$ and extending 14.4 Mm upward into the corona with a horizontal grid-size of 48 km. The model includes the effects of nonequilibrium ionization of hydrogen. The main free parameter in this class of `Bifrost` models is the initial magnetic field configuration and whether or not magnetic fields are inserted into

the computational domain through the bottom boundary. `en024048_hion` starts with two areas of opposite polarity some 8 Mm apart with an average zero signed flux. In the photosphere, the average unsigned flux is 48 G.

The two patches of opposite polarity are clearly identifiable throughout the simulation time (after the initial relaxation) of one hour. Between the opposite polarities there are loop-like structures. Proceeding from the photosphere, where plasma β is greater than one in most locations, the magnetic field expands and fills all space above heights of some $z = 1$ Mm. As the field expands and the energy density of the field becomes comparable with the thermal energy density, strong field regions are heated to higher temperatures than neighboring regions. In the model this occurs through Joule dissipation or, in the case of reconnection, through a combination of Joule dissipation and viscous heating (as plasma accelerated through reconnection becomes thermalized). However, it should be noted that the treatment of resistivity in numerical models such as `Bifrost` necessarily involves the introduction of artificial resistivity. The question then becomes, Does this treatment change the interpretation of the physics and diagnostics of the chromosphere?

Lundquist number (*S*):

$$S = \frac{L v_A}{\eta},$$

in which L is the typical length scale, v_A is the Alfvén speed, and η is the magnetic diffusivity

Reconnection and Joule dissipation are important processes when studying the heating of the chromosphere, the emergence of flux into the low atmosphere, and the generation of dynamic events as magnetic flux systems interact at various heights. In models in which resistivity is localized, such as in anomalous resistivity models (e.g., Ugai & Tsuda 1977) or in hyperdiffusion models, reconnection becomes the fast, Petschek type, and the reconnection rate is almost independent of the Lundquist number. These types of models give rise to the formation of multiple plasmoids in the reconnecting current sheet by the tearing mode instability. Observations indicate that the reconnection rate in the solar atmosphere is fast, perhaps implying that the global magnetic topology may not be sensitive to the details of the reconnection process. By contrast, the amount of energy that can be stored in an emerging flux region and the dimensions of the reconnection site itself are dependent on how reconnection proceeds, as is the redistribution of the energy released, e.g., into waves, kinetic energy, and so on (see Isobe et al. 2008). However, models including the generalized Ohm's law show that in the nearly neutral chromospheric plasma, the ambipolar dissipation rates can become as large or larger than those given by the artificial numerical rates required for numerical stability (Martínez-Sykora et al. 2012). To resolve these questions, a comparison is necessary between the observational ground truth and predictions of numerical modeling that uses various recipes for describing the microscopic physics of the chromospheric medium.

5.1.2. Comparison between models and observations. Although the Ca II 854.2-nm line, typical of the lower to mid-chromosphere, gives good agreement between observations and models, the Mg II h and k lines observed with IRIS and formed in the upper chromosphere do not.

Figure 6 shows a comparison between an IRIS observational data set and synthetic observations from the `en024048_hion` simulation. The IRIS slit goes through quiet Sun areas with some network patches and some loops seen in the Si IV slit-jaw image, whereas the synthetic slit is close to some enhanced network patches. Overall, there is a good correspondence between the synthetic intensities and the IRIS data, but both the C II lines and the Mg II k line are too narrow in the synthetic data. A comparison of the Mg II k peak intensities and the peak separation is shown in **Figure 7**. The `en024048_hion` simulation shows a broader Mg II k peak intensity distribution than it does in the IRIS quiet Sun data. This is partly because the distribution of the synthetic intensities includes the enhanced network pixels, whereas the IRIS distribution is taken from a very quiet area. However, there are also many more synthetic profiles with very faint peaks than there are in the observations. The peak separation is on average only half of what is in the observations, although the tail of the distribution extends to large peak separations.

Further insight into the nature of this discrepancy comes from optically thin lines that are formed in the chromosphere. The width of an optically thin line can place constraints on the

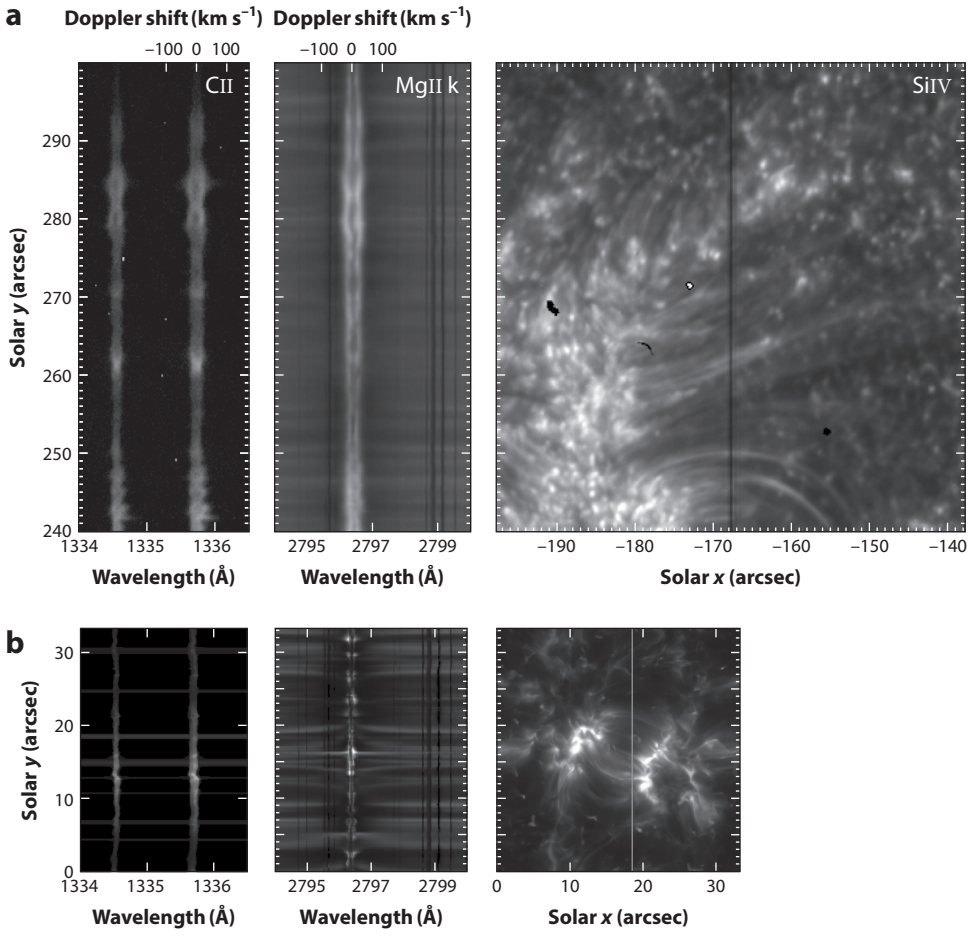

Figure 6

IRIS observations (*a*) and synthetic observations of the same lines from the `Bifrost` simulation `en024048_hion` (*b*). The spectra in the two left panels come from the location of the slit seen as a dark vertical line in the IRIS slit-jaw image (panel *a*, *far right*) and as a white line in the synthetic slit-jaw image (panel *b*, *far right*), respectively. Abbreviation: IRIS, Interface Region Imaging Spectrograph.

unresolved motions along the line of sight and determine how much of the discrepancy in the peak separation is caused by such motions (see Section 3). One such optically thin line formed in the chromosphere is the OI line observable with IRIS at 135.56 nm (Lin & Carlsson 2015; see also Section 3). The amount of nonthermal broadening is constrained by observations to less than 7 km s^{-1}, which is much less than what would be needed to explain the observed widths of the MgII h and k lines if the observed widths of the MgII h and k lines were caused by nonthermal broadening.

For an optically thick line, the gradient of the source function with height is also very important, giving rise to what is often called opacity broadening (see Rathore et al. 2015a). Too narrow lines in the simulations thus tell us that the local peak of the source function (giving rise to the k_2 intensity peaks) is at a smaller column mass than it is in the observations. This could be caused by too little heating in the low chromosphere in the simulations.

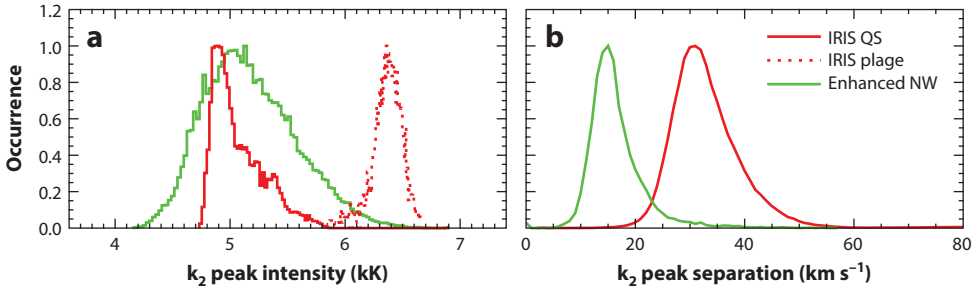

Figure 7

IRIS observations (*red*) and synthetic observations from the `Bifrost` simulation (*green*). (*a*) Radiation temperature of the Mgii k_2 peak intensity and (*b*) the Mgii k_2 peak separation. Abbreviations: IRIS, Interface Region Imaging Spectrograph; NW, network; QS, quiet Sun.

In quiet Sun areas, the Oi 135.56-nm line itself is also observed to be wider than what we find in the simulations (Lin & Carlsson 2015), indicating that there are not enough small-scale dynamics in current models. This could be due to these models being run at too low spatial resolution. Simulations run at higher spatial resolution (31 km horizontal grid spacing instead of 48 km) show increased velocity dispersion in the chromosphere (but still not enough to match the line widths).

What is the cause of this discrepancy between the state-of-the-art models and observations? We have already mentioned that increased numerical resolution in the models causes more dynamics in the chromosphere and therefore more broadening and some heating by dissipation of waves. Another possible explanation comes from the way the initial magnetic field configuration is set up in the `Bifrost` models, through specifying the field at the bottom boundary with a potential field extrapolation filling the rest of the volume. This means that there is no way to initialize the simulation with magnetic field variations at small scales in the photosphere, because structure in a potential field decays with a scale height equal to the spatial scale. Furthermore, the large scales are inserted directly into the simulation box instead of emerging through the atmosphere. Flux emergence would lead to mass being brought up with the field and to lower-lying, mass-loaded field lines. Both effects would lead to stronger magnetic fields in the low and middle chromosphere than there are in the standard model and more magnetic heating. A third possibility is missing physics in the simulation. Ion–neutral effects certainly play a role in the low and middle chromosphere, providing heating through ambipolar diffusion (Section 2.4), concentrating currents, and changing the way magnetic field emerges into higher layers (Martínez-Sykora et al. 2017a). Other plasma physics effects could also play a role (Section 5.2).

5.2. Plage

Active regions have much stronger average magnetic field than quiet Sun. In regions in which the magnetic field is strongest, sunspots occur, because of the reduced or suppressed magnetoconvection. Much of the rest of the active region is covered by plage, which comprises regions of enhanced chromospheric brightness, and by regions that appear quiet-Sun-like in the photosphere but that are usually covered (in the chromospheric passbands) by strong canopy fields connecting regions of opposite polarity. The chromosphere is brightest in plage regions, which are at the footpoints of most coronal loops and thus act as an important interface through which the nonthermal energy driving coronal heating must pass. Plage also provides a significant fraction of the UV emission that impacts the Earth's upper atmosphere. Yet, the properties of plage and what causes its heating remain poorly understood.

The advent of high-resolution spectra in a variety of spectral lines that are sensitive to chromospheric conditions has led to some recent progress in better constraining the properties of plage. In plage, IRIS Mg II 279.6-nm spectral line profiles are most often singly peaked, avoiding the central reversal (k_3 feature) that dominates such profiles in most other regions on the Sun. This type of profile cannot be explained by the FAL P model (Fontenla et al. 1991) and seems to provide quite stringent constraints on the temperature stratification. Preliminary semiempirical forward modeling suggests that these profiles can be caused by a steep temperature increase in the low chromosphere, followed by a relatively flat (with height) temperature plateau all the way up to the TR. The latter has to occur at high column mass to ensure coupling of the source function to the local conditions (Carlsson et al. 2015). This is compatible with recent work based on the novel STiC inversion code (de la Cruz Rodríguez et al. 2016), which also shows that plage is characterized by an increased microturbulence that is remarkably spatially homogeneous over the plage region (also supported by the nonthermal broadening in O I 135.6 nm; Carlsson et al. 2015). The relative spatial uniformity may be related to the expansion with height of the magnetic field, which is highly structured in the photosphere, into a volume filling field with canopy fields in between photospheric flux concentrations. This is supported by observations indicating significant chromospheric heating above and in the close vicinity of strong photospheric flux concentrations, as evidenced by the reversed core profiles in the Ca II 8542-Å profiles (de la Cruz Rodríguez et al. 2013). However, the spatiotemporal properties of these profiles suggest that the heating associated with the magnetic field is quite structured and variable in time. Spectropolarimetric observations are key to further investigating the magnetic field in plage and how it impacts the local thermodynamics.

A wide range of heating mechanisms for plage regions have been proposed. Magnetoacoustic shocks are plentiful in plage and appear to dominate the dynamics of the middle and upper chromosphere, with the generation of dynamic fibrils (see Section 5.4). However, it is not fully clear how important these shocks are for the energy balance of plage, with some results suggesting a substantial contribution (e.g., Rezaei et al. 2007, Sobotka et al. 2014) and others suggesting a significant magnetic-field-dependent component (Beck et al. 2013b, Barczynski et al. 2018). The magnetic-field-dependent component could be caused by a variety of mechanisms that are related either to dissipation of magnetic waves or to heating associated with magnetic reconnection.

Although observations of wave dissipation are scarce (see Section 5.5), Alfvén waves carrying a significant energy flux have been observed throughout the chromosphere, first in spicules (De Pontieu et al. 2007c) and later in various other chromospheric features. There is also evidence for a variety of other wave modes [for a review, see Jess et al. (2015)]. Several theoretical models invoke transverse waves to drive heating in magnetic network or plage regions. For example, high-frequency (periods less than 100 s) magnetoacoustic waves triggered either by nonlinear coupling to upward propagating transverse waves generated in the photosphere (Hasan & van Ballegooijen 2008) or by reconnection between weak granular fields and strong flux concentrations in the photosphere (Isobe et al. 2008) have been shown to lead to significant heating in the chromosphere, but observational studies examining whether such waves are present at sufficient power are needed to address this proposed heating mechanism. Strong heating could also come from Alfvén waves, generated, for example, through the small-scale interaction of magnetic flux concentrations with the convective flowfield and dissipated through the turbulence resulting from nonlinear interactions with counter-propagating waves reflected off the TR (van Ballegooijen et al. 2011). However, it is unclear how a more realistic treatment of the chromospheric propagation of these waves (and their potential nonlinear coupling to slow-mode shocks, e.g., Brady & Arber 2016) would affect the proposed heating rate.

Magnetic reconnection is also a major candidate for heating plage. It seems that large-scale flux emergence (on supergranular scales) leading to reconnection from subsequent cancellation of opposite polarity fields and violent heating does not have a large enough spatiotemporal filling factor to produce the emission in quiescent plage (Leenaarts et al. 2018). However, the interaction between granular-scale weak fields (which are quite common in plage; Buehler et al. 2015) and strong plage fields could in principle lead to the heating in quiescent plage. It is not yet fully clear how pervasive the mixed polarity weak fields are in the vicinity of the mostly unipolar plage regions (because of their weak signal in polarimetric observations). However, the existence, if confirmed, of small-scale coronal loops associated with plage regions (in particular, those associated with high-temperature upper TR plasma or moss; Berger et al. 1999), as deduced from intriguing high-resolution coronal observations (e.g., Wang 2016, Barczynski et al. 2017), warrants further investigation of this idea (Yurchyshyn et al. 2010). Another possible heating mechanism in plage is dissipation of currents resulting from braiding. This could be similar to the Joule dissipation modeled in `Bifrost` simulations of quiet Sun enhanced network. Alternatively, perhaps the current generation and dissipation are similar to what is found in 2.5D MHD simulations of spicule formation that include ion–neutral interactions (Martínez-Sykora et al. 2017b)? Evidence for such currents, their dissipation, and the associated heating of plasma has been deduced from IRIS observations of TR network jets (De Pontieu et al. 2017a). Perhaps these currents extend into the chromosphere and are dissipated by ambipolar diffusion? Such a heating scenario would be compatible with earlier suggestions that spicules or "straws" play a substantial role in creating the haze of emission around strong network and plage regions (Cauzzi et al. 2009).

To address all these questions, quasi-realistic numerical models that incorporate some of these physical mechanisms are urgently needed. However, there currently are no such models. Experiments with `Bifrost` have so far failed to produce a convincing plage chromosphere. Apparently, introducing a strong average magnetic field in the numerical domain is not sufficient to create plage. Simulations with unipolar magnetic field with an average unsigned field strength of order 200 G (the canonically reported values for plage) but without opposite polarity or granular-scale mixed polarity fields lead to an atmosphere that is more like that of a coronal hole (i.e., a quiet Sun chromosphere and very low-temperature corona). The apparent necessity (for the formation of strong plage) of an opposite polarity plage patch and a hot, dense corona connecting the two polarities would fit with the observed correlation between the strongest plage emission and the presence of upper TR moss (which occurs at the footpoints of the hottest and densest loops) (Carlsson et al. 2015). Numerical simulations with a numerical domain large enough to include two very strong magnetic field patches are a critical first step toward making progress. To properly test the various proposed heating mechanisms, further generations of such a model would have to include ambipolar diffusion, nonequilibrium ionization, small-scale mixed polarity fields, and higher spatial resolution (e.g., to allow more efficient driving of waves). Looking further into the future, it is also important to strive for multifluid simulations that consider the complex interactions between the various fluids (ionized species, neutrals, and electrons) in the chromosphere. These could capture non-MHD effects that have been proposed as potential heating mechanisms, e.g., the Farley–Buneman instability caused by drift between species (Fontenla 2005, Gogoberidze et al. 2014, Fletcher et al. 2018) or electrical currents generated by ion-neutral drag (Krasnoselskikh et al. 2010). So far, non-MHD effects have only been tested in static, semiempirical background atmospheres, which do not accurately capture the nonequilibrium ionization on which these effects critically depend. Multifluid simulations will also provide key insights into the FIP effect.

Until such models become available, observations should focus on key constraints that will help guide model development. For example, careful comparisons of the heating associated with shocks (e.g., from ALMA), assisted by observations of shock signatures (sawtooth patterns in

spectral lines; Section 5.1), and determination of the spatiotemporal filling factor of shocks can constrain their contribution to the plage energy budget. Novel approaches using photospheric data to determine the Poynting flux (Welsch 2015) can also help constrain chromospheric heating mechanisms, especially when compared with good proxies for chromospheric heating (e.g., from ALMA or inversions of chromospheric lines). Spectropolarimetry with unprecedented sensitivity is expected from DKIST observations. Coupled with state-of-the-art inversion techniques and high-cadence spectroheliograms (e.g., using integral field units to capture the highly dynamic nature of chromospheric heating), these observations can help clarify not only the contributions of mixed polarity fields in heating plage and driving dynamics (e.g., spicules) but also the role of current dissipation in the plage chromosphere. By providing velocities and magnetic field at various "heights" in the chromosphere, such data can also be used to properly determine waves and wave modes, and, possibly, to study wave dissipation (see Section 5.4).

A better understanding of what heats the plage chromosphere will likely lead to new insights into coronal heating. Not only is the plage chromosphere the conduit for all nonthermal energy that reaches the corona but also there are tantalizing indications for a connection between chromospheric and coronal heating, as suggested by observations that the strongest plage emission arises in the same regions where the upper TR moss emission is brightest and, thus, the coronal pressure is highest (Carlsson et al. 2015). More observations and numerical modeling are needed to determine the nature of this connection.

5.3. Flux Emergence

Flux emergence is a key process that plays a role in filling the chromosphere and corona with magnetic field and at the same time contributes directly to the heating of the chromosphere, even if its role in the average state of plage or network remains unknown. We know that large-scale emergence locally has a large heating effect but likely does not have the spatiotemporal filling factor to explain the average state. Just how important granular-scale flux emergence is remains unclear. Note that for a substantial fraction of the active region lifetime, except during immediate emergence, there is little large-scale emergence going on, yet the chromosphere is bright anyway. Thus, the emergence of field at active region size scales presumably does not play a direct role during most of the active regions' lifetime. The bulk of the heating must therefore be sought from other sources, as described in the plage section above. Nevertheless, flux emergence is obviously a vital process in filling the corona with magnetic field, a process that is accompanied by violent reconnection observable in diagnostics formed throughout the entire span of the chromosphere as suggested theoretically and by simulations (Heyvaerts et al. 1977, Galsgaard et al. 2007, Archontis & Hansteen 2014, Chatterjee et al. 2016, Hansteen et al. 2017).

A current paradigm of flux emergence is that flux tubes are transported by buoyancy, with the strongest eventually forming Ω-shaped loops that rise and emerge through the surface to form sunspots (Zwaan 1987, Moreno-Insertis 2007, Fan 2009). The amount of twist in emerging flux tubes has important consequences for the fate of the rising flux (e.g., see Cheung & Isobe 2014). The outer layers of the Sun are extremely stratified, with a density contrast of some 10^6 from the bottom of the convection zone to the photosphere and even more so between the photosphere and corona, where a density contrast of 10^8 is found (see **Figure 2**). In the photosphere and above, the atmosphere is no longer convectively unstable, and the rise of a flux tube is halted, at least temporarily.

Field stored just below or rising to the photosphere will break through the surface and enter the upper atmosphere once the gradient of the field strength becomes sufficiently large (Acheson 1979). The instability mode responsible for the breakthrough is most likely the undular mode in which the perturbation vector is parallel to the magnetic field. In this mode plasma falls along

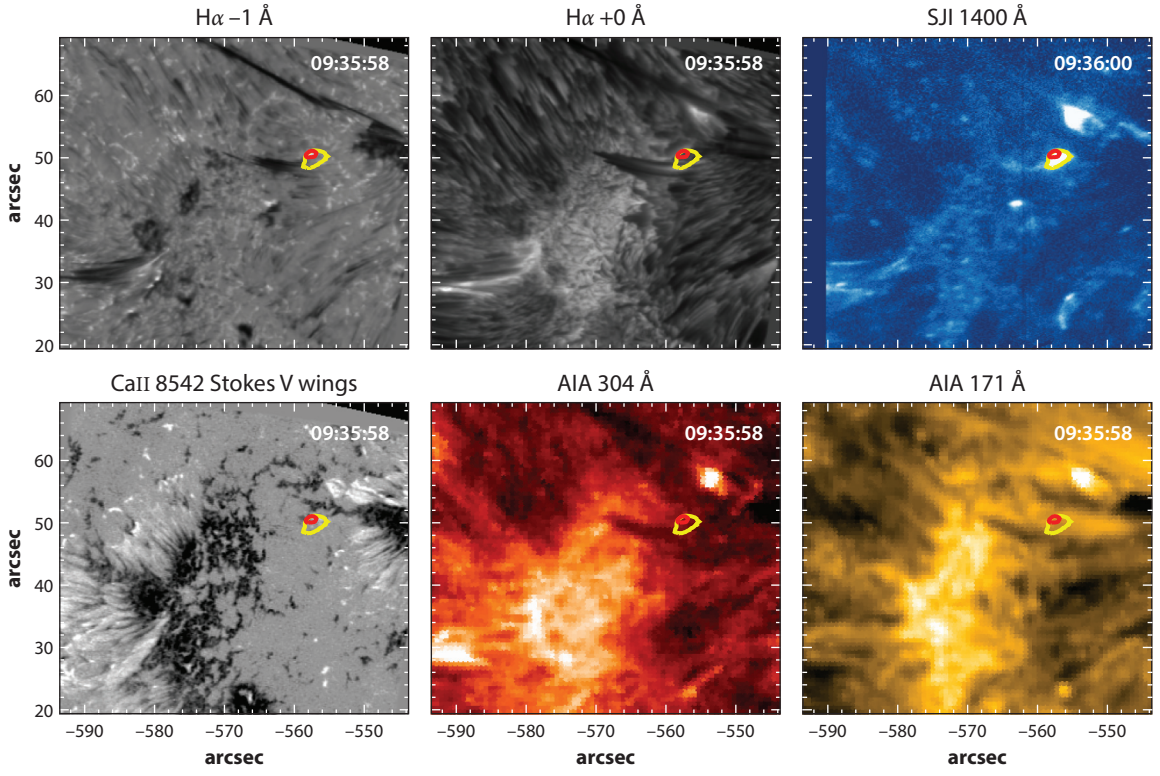

Figure 8

Example of reconnection resulting from emerging flux and leading to Ellerman bombs (*red contour*) and UV bursts (*yellow contour*) as seen in a variety of diagnostics. Abbreviations: AIA, Atmospheric Imaging Assembly; SJI, slit-jaw image.

the field lines so that rising crests get lighter while troughs get heavier, forming a set of so-called U-loops. The opposite polarity legs of such U-loops can become pinched in intergranular lanes, leading to magnetic reconnection and the formation of plasmoids (Tortosa-Andreu & Moreno-Insertis 2009). The implied magnetic structure from repeated reconnection events near the photosphere is therefore that of serpentine field lines with undulations of wavelength 2 Mm (the wavelength of maximum growth rate of the undular mode; Pariat et al. 2004); reconnection will occur in the magnetic dips, sometimes forming EBs (Hansteen et al. 2017), releasing denser material to the solar interior while allowing the further emergence of magnetic flux into the chromosphere and corona (Lites 2009, Centeno 2012). Ortiz et al. (2014) find that flux emerging through the chromosphere incurs the formation of cool bubbles, expanding horizontally and ascending with a speed of some 5 km s^{-1}. As the bubbles expand they will at some point interact with each other or with the preexisting ambient field. Such locations are the sites of large-angle reconnection at all levels of the chromosphere.

An example of such large-angle reconnection lies in EBs and UV bursts (UVBs) observable in newly formed active regions where magnetic flux is still vigorously emerging. In **Figure 8**, we see such a region containing an EB and a UVB in the same location: In the EB, the wing of Hα as well as the wing of the CaII triplet line at 854.2 nm are bright while the line cores remain dark. At the same time, we find a brightening of the MgII and SiIV slit-jaw images and 170-nm images from the Atmospheric Imaging Assembly (AIA; Lemen et al. 2012) onboard the *Solar Dynamics Observatory* (SDO). Gas is being energized for the entire temperature range spanned by the chromosphere.

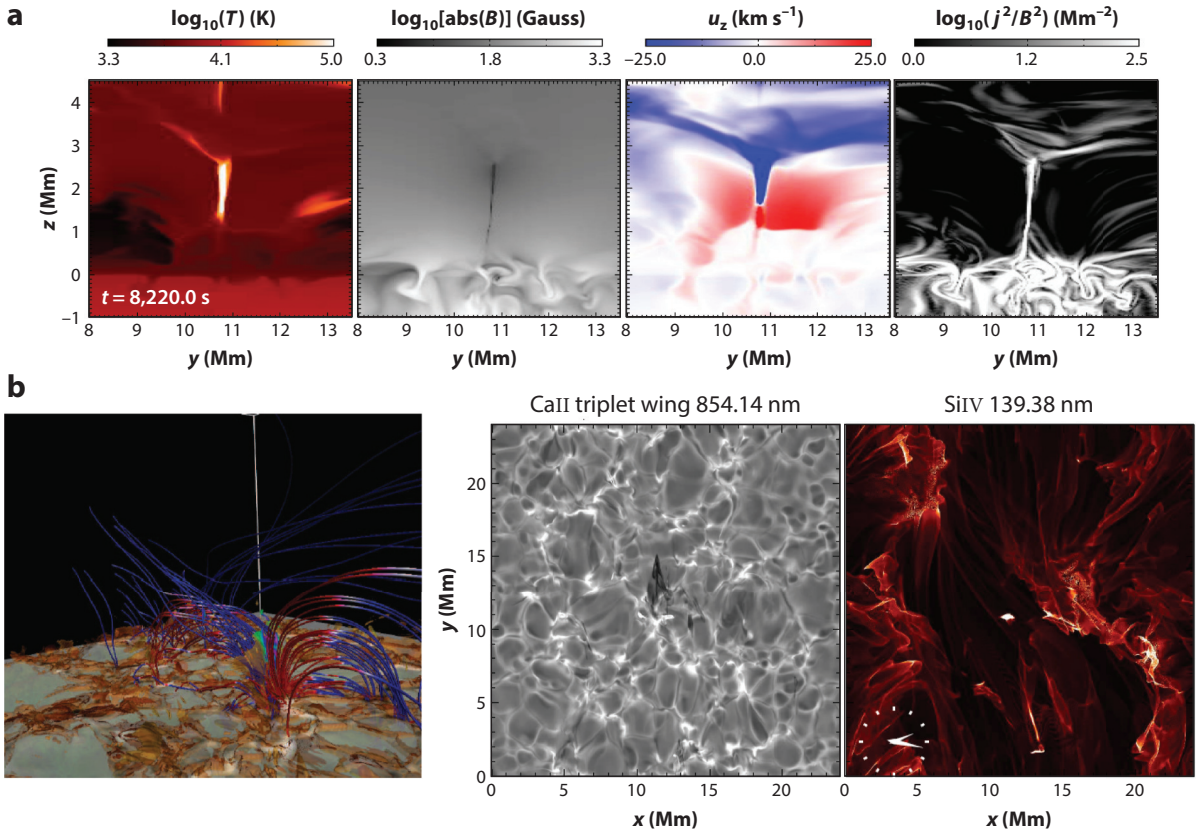

Figure 9

Numerical simulation of flux emergence resulting in a UV burst. (*a*) From left to right: temperature, magnetic field strength, vertical velocity, and j^2/B^2, which is enhanced in the current sheet where reconnection occurs. (*b*) Magnetic field lines (*red/blue* indicate *negative/positive sign* of B_z) and Joule heating (isocontours) (*left*), intensity in blue wing of the CaII 854.2 nm triplet line (*middle*), and total intensity of the SiIV 139.4-nm transition region line (*right*).

This brightening occurs above a region of opposite magnetic polarities interacting as they are being pushed together. Although no signal is seen in the hot AIA 30.4- (0.1 MK) and 17.1-nm (1 MK) channels, a dark surge is visible in both Hα and CaII as well as with AIA: Cool gas of considerable opacity is injected to great heights in the chromosphere.

This is borne out by numerical modeling of flux emergence. In **Figure 9**, we show the current sheet formed in the chromosphere as a result of cool bubble interaction as opposite polarity field is pushed together, initiating reconnection and vigorous heating of the plasma. This heated gas is covered by a canopy of cooler gas brought up earlier as part of the flux emergence process that has not yet had time to drain back to the photosphere. Bidirectional jets, with velocities reaching several hundred kilometers per second, are generated, expelling both hot and cooler gas upward toward the corona and down toward and through the photosphere. Synthetic observables of the blue wing of the CaII 854.2-nm line (some 25 km s^{-1} from line center) show both brightening, similar in structure to an EB, and a surge. The modeled SiIV 139.4-nm line formed at TR temperatures has extremely increased emission, at least two orders of magnitude brighter than the average, which is a typical signal of UVBs.

This example shows that the active chromosphere, during periods of active emergence, can be quite complex in structure and is characterized by velocities more typical of the Alfvén speed ($\sim$100 km s^{-1}) than the speed of sound ($\sim$10 km s^{-1}) and by extreme gradients in temperature and density and is where cool plasma often overlies hot plasma. Is then the magnetically dominated chromosphere energized by such events of varying sizes? When considering the structure of CaII K line emission in the chromosphere during a period of active emergence, Leenaarts et al. (2018) concluded that the bulk of radiative losses and hence heating in this emerging region is gentle, originating in more spatially extended regions correlated with sites of high horizontal field strength and, rather than being episodic, is persistent, even during active emergence. This is also evident in **Figure 8** in regions far from the EB/UVB and for active regions where new field emergence has ceased.

Small-scale emergence could potentially play an important role in the energization of the chromosphere, both in active regions and in the quiet Sun. The arrival of emerging field in the photosphere at granular scales has been observed by several authors, starting with tantalizing glimpses seen in ground-based data by De Pontieu (2002). The launch of the *Hinode* satellite and the availability of seeing-free magnetograms led to the realization that the Sun is pervaded by continuously emerging small-scale horizontal field (Centeno et al. 2007, Ishikawa et al. 2008, Ishikawa & Tsuneta 2009). Some of this flux is likely due to a local dynamo (seeded by field brought up earlier and stored below the photosphere?) rather than due to the direct emergence of large-scale magnetic structures from below (see Rempel 2014). Observations by Martínez González et al. (2007) and Martínez González & Bellot Rubio (2009) show that a significant proportion of this field likely reaches the chromosphere. Estimates of the energy flux provided to the chromosphere by rising flux is of order 1–10 kW m^{-2}, which is sufficient to be of interest to the chromospheric and coronal energy budget. The questions remain whether sufficient field actually reaches the chromosphere and, if so, whether heating mainly occurs through large-angle violent reconnection or through Joule dissipation as fields are brought together at smaller angles. Gošić et al. (2018) studied small-scale ($\sim$1 Mm or less) brightenings associated with flux cancellation in the internetwork magnetic field as observed in the MgII h and k lines. They conclude that cancellations produce clear signatures of heating in the upper atmospheric layers. However, at the resolution and sensitivity accessible to the authors using SST and IRIS, their number density appear to be one order of magnitude too low to account for global chromospheric heating. Weak fields, nonobservable by current spectropolarimeters, could, in principle, supply the shortfall as they ascend to the chromosphere, heating plasma through braiding or reconnection with the preexisting field.

5.4. Spicules

Jets are a common phenomenon in the solar atmosphere, from X-ray jets at the largest scales down to chromospheric spicules on the subarcsecond scale. Jets are of broad interest: (*a*) Their formation has remained poorly understood; (*b*) jets have been implicated in playing a substantial role in the energy and mass balance of the chromosphere, TR, and corona; and (*c*) they are common in other astrophysical contexts.

Spicules are the most ubiquitous of all jets, as evidenced by the fingerlike protrusions (up to 15 Mm above the limb) that dominate the visual appearance of the solar limb in chromospheric lines like Hα (see **Figure 1**). Despite a multitude of studies in the second half of the twentieth century (Beckers 1968), many of their properties remained controversial, including their disk counterparts, in large part because previous generations of instrumentation were not sufficient to capture their evolution, which occurs on subarcsecond spatial scales and timescales of tens of seconds and shorter. This led to a plethora of many competing theoretical models, so that their formation and impact on the outer atmosphere remained mostly shrouded in mystery (Sterling

2000). During the past decade or so, the arrival of high-resolution instrumentation and advanced numerical simulations has led to significant breakthroughs in our understanding of spicules.

High-quality time series of Hα filtergrams using the SOUP filter at the SST revealed the evolution of short, highly dynamic fibrils in and around active region plage when observed on the disk (Hansteen et al. 2006, De Pontieu et al. 2007a). The impeccable stability of the ground-based time series, enabled by powerful adaptive optics, for the first time showed how these features (dynamic fibrils) are characterized by parabolic paths in space–time plots, with up- and downward motions of order 10–40 km s^{-1}, with striking correlations between acceleration and maximum upward flows, lifetimes of order 3–5 min, and quasi-periodicities of 3–5 min. Comparisons with advanced radiative MHD models showed remarkable similarities, suggesting that these features are driven by slow-mode magnetoacoustic shocks that propagate upward along the strong magnetic field in the flux concentrations that constitute plage regions. Radiative MHD simulations suggest that these relatively weak shocks can be driven by several essentially acoustic mechanisms, including leakage into the chromosphere of (*a*) convective motions; (*b*) photospheric p-mode oscillations along inclined magnetic field lines (Michalitsanos 1973, Bel & Leroy 1977, Suematsu 1990), which had previously been shown to produce 5-min oscillations in the dynamic fibril formation (De Pontieu et al. 2004); and (*c*) pressure pulses associated with gentle heating from current dissipation in the photosphere. Evidence for a similar mechanism has been seen in quiet Sun mottles on the disk [Rouppe van der Voort et al. (2007), and see Suematsu et al. (1995) for earlier work]. However, these slow-mode shocks are driven by thermal pressure perturbations and, thus, remain relatively weak and lead to chromospheric protrusions into the corona that reach heights of only a few megameters in recent models that include realistic radiative losses (Heggland et al. 2007, Martínez-Sykora et al. 2009). At the limb, they are only visible above active region plage (and for historical reasons often called type I spicules, see below), where their height is modest (typically only 5–6 Mm; Pereira et al. 2012), which is compatible with historical reports that spicules are suppressed above plage (Beckers 1968). The presence of these shocks in plage, at the footpoints of many active region coronal loops, has a significant impact on the above-lying active region TR and corona (discussed in Section 6).

In the weaker magnetic field environment of quiet Sun and coronal holes, the flux tubes expand more rapidly with height, diluting the shock wave energy and producing only very minor excursions into the corona that are not clearly visible at the quiet Sun or coronal hole limb, given the line-of-sight confusion that favors taller features. In summary, these shocks, driven by essentially acoustic pressure perturbations, thus do not seem to lead to limb spicules in much of the Sun. This conclusion matches earlier common wisdom, which had rejected the slow-mode shock scenario as a viable mechanism (Sterling 2000) for producing tall spicules after it was first explored in idealized ad hoc simulations (e.g., Hollweg 1982).

Confusingly, the properties of the shock-driven on-disk dynamic fibrils are similar to what low-resolution twentieth-century observations reported for most of the quiet Sun and coronal hole limb (10–40 km s^{-1}, up- and downward motions, 5-min lifetimes). The confusion was mixed with shock and awe when observers got their first glimpse (De Pontieu et al. 2007b) of seeing-free CaII H filtergram time series of quiet Sun and coronal hole limb spicules obtained with the *Hinode*/SOT at very high cadence (1.6 s) and resolution (0.16 arcsec), which showed much faster spicules (30–100 km s^{-1}) that have much shorter lifetimes (of order 1–2 min) than the canonical values from previous reports (Beckers 1972), with mostly upward motions followed by rapid fading. Because of these discrepancies with previous observations, these features, which dominate the quiet Sun and coronal hole limb, were called type II spicules (De Pontieu et al. 2007b). Some of this confusion may have been resolved by Pereira et al. (2013b), who showed that the poor spatiotemporal resolution of pre-*Hinode* observations masks the very high upward speeds of type II

spicules and can give the appearance of up- and downward motions leading to the lower speeds and longer lifetimes previously reported.

There has been vigorous debate in the community concerning whether there really are two types of spicules, with some observers claiming that CaII H spicules typically show up- and downward motions (Zhang et al. 2012) at low speeds, whereas analysis of the same data set with other techniques reveals mostly upward motions at high speeds (Pereira et al. 2012). A more unified picture has appeared from combined *Hinode* and IRIS observations, which confirmed the high speeds in type II spicules (Pereira et al. 2014, Skogsrud et al. 2015) and revealed qualitatively different behavior of both types of spicules in the above-lying TR. On the one hand, type I spicules (in active region plage) are accompanied at their leading edge or top by compact brightenings in TR lines (e.g., SiIV 140.3 nm) (Skogsrud et al. 2016). Such brightenings are compatible with a scenario in which slow-mode magnetoacoustic shocks compress the plasma as they reach the TR, which is driven upward by the cool spicular flows in the wake of the shock. Type II spicules, on the other hand, appear to show TR emission along their whole length, often rapidly appearing at high propagation speeds when observed at the limb (Pereira et al. 2014, Skogsrud et al. 2015) and on the disk (Tian et al. 2014, Rouppe van der Voort et al. 2015, Henriques et al. 2016). The IRIS observations also show that many type II spicules can be faintly detected in the MgII h and k lines during both the upward and downward phases, in contrast to the CaII H, where such detections are quite rare and the upward phase dominates. This suggests a scenario in which the opacity of CaII H type II spicules rapidly drops during the spicule lifetime, either because of rapid temperature changes (as suggested by the appearance of rapid heating from IRIS SiIV observations) or because of density changes [supported by the fact that MgII k, which has 18 times higher abundance than CaII H, often shows spicules during the downward phase, and this is compatible with findings by Beck et al. (2016)].

The fundamental differences between type I and type II spicules are also present when observed on the disk. The disk counterparts of type I spicules are visible as dynamic features that grow and shrink over the course of several minutes in the inner wings ($\approx$20–30 km s^{-1}) of chromospheric lines such as Hα (and sawtooth patterns in λ-time plots). The disk counterparts of the more common type II spicules present as very short-lived ($\approx$30 s) excursions in the far blue wing (40–80 km s^{-1}, caused in part by strong upflows) of chromospheric lines (Rouppe van der Voort et al. 2009; Sekse et al. 2012, 2013a). Observations close to the limb show that these rapid blueshifted excursions or RBEs smoothly transition from absorbing features on the disk into emission features in spicules off the limb, establishing their identity as the long-sought-after disk counterpart of limb spicules. Detailed studies of these RBEs show that they most often rapidly (with speeds of more than 50 km s^{-1}) propagate away from the close vicinity of network and plage flux concentrations while undergoing significant swaying and torsional or azimuthal motions associated with transverse MHD and Alfvénic waves (Sekse et al. 2013a,b; Kuridze et al. 2015), with the latter leading to rapid excursions in the red wing (rapid redshifted excursions or RREs) near RBEs. Such wave motions are also evident from limb observations of type II spicules (De Pontieu et al. 2007c; Okamoto & De Pontieu 2011; De Pontieu et al. 2012, 2014a).

These new observations have provided unprecedented constraints for theoretical models of spicules. Although the type I spicules naturally form in the self-consistent `Bifrost` simulations (which only depend on the magnetic field as a free parameter), the much faster type II spicules had been mostly absent, with only a single event captured in many solar hours of simulations (Martínez-Sykora et al. 2011). This puzzling discrepancy motivated studies of ion–neutral interaction effects, which recently have led to a breakthrough in our understanding of how type II spicules can form. Advanced numerical simulations show that ambipolar diffusion can fundamentally change the interaction between the strong network or plage magnetic fields and

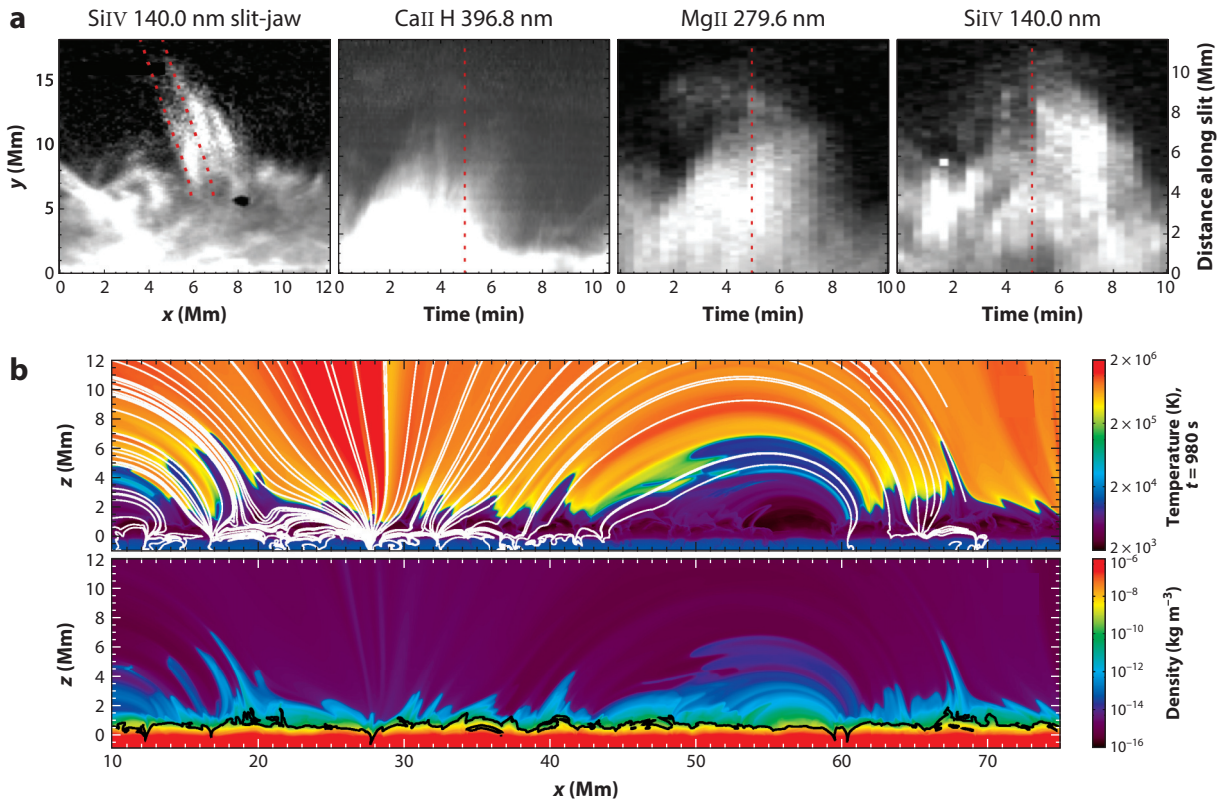

Figure 10

Observations and numerical models of spicules. (*a*) Thermal evolution of type II spicules (*far left*, SiIV 140.0-nm IRIS slit-jaw image) suggests heating of at least a subset of spicule threads from chromospheric temperatures (*Hinode* CaII H 396.8-nm space-time plot) to upper chromospheric (IRIS MgII 279.6-nm) and transition region (SiIV 140.0-nm) temperatures. (*b*) State-of-the-art numerical models show cool and fast intrusions of chromospheric plasma that resemble type II spicules rapidly propagating into the tenuous hot corona. Panel *a* adapted from Skogsrud et al. (2015), and panel *b* adapted from Martínez-Sykora et al. (2017b). Abbreviation: IRIS, Interface Region Imaging Spectrograph.

the ubiquitous granular-scale weak fields (Martínez-Sykora et al. 2017b). It allows these tangled weak fields to diffuse into the upper chromosphere where the violent release of magnetic tension (typically not associated with magnetic reconnection) can lead to rapid acceleration of plasma to drive supersonic jets with speeds of 50–100 km s^{-1} that are heated through the diffusion of ambipolar currents. The release of magnetic tension also naturally leads to the generation of transverse magnetic waves, providing a neat explanation for the observed waves. The jets are not present in previous advanced numerical simulations and only appear when ambipolar diffusion is included (**Figure 10**). Some of the jets appear to be driven by a second mechanism involving wave-mode coupling around the $\beta = 1$ surface. Synthetic observables from these simulations show remarkable agreement with chromospheric observations of spicules from SST and IRIS, including the significant increase in emission of TR lines observed with IRIS and *Hinode*/SOT (Skogsrud et al. 2015). These results are based on 2.5D radiative MHD simulations assuming equilibrium ionization and including the generalized Ohm's law to approximate the multifluid interactions between ions and neutrals. However, nonequilibrium ionization of hydrogen and helium can drastically change the expected ionization degree (and thus amount of ambipolar diffusion). In addition, the

assumptions underlying the generalized Ohm's law can break down when the decoupling between ions and neutrals becomes too large, which may occur at the top of the chromosphere. To help understand whether this mechanism is indeed a dominant formation mechanism of type II spicules, future simulations will need to include three dimensions, nonequilibrium ionization, and possibly multifluid treatment.

Spicules can likely be formed through a variety of mechanisms. The challenge is to determine which of the proposed mechanisms is the most common driver. Some of the alternative mechanisms for spicule formation that have been proposed are shocks driven by strong vorticity in the photosphere (Iijima & Yokoyama 2017) and the eruption of mini-filaments that form close to neutral lines (Sterling & Moore 2016). The former model is based on numerical simulations but with the identification of spicules based on the comparison of the physical simulation variables (e.g., upward velocities) with those derived from observations. However, such a comparison can be misleading and does not capture the complex thermodynamic constraints provided by the detailed chromospheric and TR spectral line profiles. To determine whether spicules are often driven by this mechanism, such a detailed comparison is critical. This is illustrated by Martínez-Sykora et al. (2013), who identified spicules based on physical variables but found, upon further investigation, that they could not reproduce the disk observations of spicules. Similarly, it will remain unclear how important the mini-filaments eruption mechanism is for spicule formation until a numerical model fleshes out the current cartoon model.

Until such advances in the numerical models are made, observations can meanwhile provide only stringent tests of some of the predictions of the currently proposed mechanisms. Such tests are best performed using disk-based data, because they allow correlations between the photospheric roots and the formation of chromospheric spicules. This can be done, for example, through detailed studies of the dynamic correlations between the weak, granular-scale fields and launching of RBEs (which would test the model by Martínez-Sykora et al. 2017b). Similarly, studies of correlations between photospheric vorticity and RBE formation can test the Iijima & Yokoyama (2017) model while high-resolution magnetograms can be searched for the formation of neutral lines and mini-filaments and their correlations with RBEs. The next few years thus provide exciting opportunities to gain a better understanding of the formation of spicules.

5.5. Waves

A wide variety of wave modes has been observed in the chromosphere, including kink mode waves, Alfvén waves, sausage waves, and shock waves. We do not aim to provide a review here of this extensive literature but refer the reader to the review by Jess et al. (2015). These waves have been utilized for seismological applications, an application that is highly challenging (except perhaps in the strongest magnetic field of sunspot umbrae) and should be undertaken with extreme care, because of the highly structured and dynamic chromospheric environment straddling the $\beta = 1$ surface, which often leads to counterintuitive interpretation of data because of the NLTE nature of many of the diagnostics.

Given these complexities, the main interest in waves lies in their potential role in heating the chromosphere and outer atmosphere or solar wind. Many of these waves have been shown to carry substantial energy fluxes into and through the chromosphere and beyond (starting with the unambiguous discovery of ubiquitous Alfvénic waves on spicules; De Pontieu et al. 2007c). However, a key ingredient to any wave heating model is the dissipation mechanism that transforms the wave energy into heat. Recent results have now provided new insights into resonant absorption, a physical mechanism (based on mode coupling of transverse magnetic waves into azimuthal modes), which has long been proposed as the dominant heating mechanism of the corona (see, e.g., Ionson 1978). Observational evidence for this process had been elusive until groundbreaking studies by

Okamoto et al. (2015) and Antolin et al. (2015), who exploited the thermal coverage and high resolution of coordinated *Hinode*/SOT and IRIS observations of prominences. Comparison with numerical models revealed telltale signatures of resonant absorption and subsequent dissipation and heating of plasma through the Kelvin–Helmholtz instability (KHI). In the model, the KHI creates small-scale vortical flows, rapidly driving the wave energy to spatial scales that are small enough to provide substantial dissipation. These vortical flows occur on small enough spatial scales that they lead to increased nonthermal broadening. Given the ubiquity of transverse waves in the chromosphere, one wonders how common this process is. Does it only occur in prominences or does it also affect other chromospheric regions? Recent results suggest that in at least a subset of spicules a similar process may be at work (Kuridze et al. 2016, Antolin et al. 2018). Future studies will hopefully clarify to what extent these TWIKH (transverse wave–induced Kelvin–Helmholtz) rolls are responsible for the pervasive and currently unexplained nonthermal broadening in plage regions (Carlsson et al. 2015), spicules (De Pontieu et al. 2014a), and the TR (De Pontieu et al. 2015).

6. CONNECTIONS BETWEEN CHROMOSPHERE AND TRANSITION REGION/CORONA

Previous generations of imagers of the TR and corona showed little indication of significant dynamic connections between these hot regions of the solar atmosphere and the underlying cool chromosphere, except perhaps during eruptions or large-scale jets. As a result, the TR and coronal dynamics and energetics have often been studied in isolation from the underlying atmosphere. The latter is often only acknowledged as a passive reservoir of mass whose properties closely mimic that of the semiempirical VAL atmosphere (Section 4.1). Although this approach is still common in studies of coronal heating, it is no longer backed up by observations. The arrival of subarcsecond resolution in imaging and spectroscopic observations of the TR (e.g., with IRIS) and corona (e.g., the HiC sounding rocket; Kobayashi et al. 2014) has resolved the subgranular spatial scales on which the intimate connections between the dense chromosphere and the above-lying hotter parts of the atmosphere become apparent.

Many phenomena that were previously thought of as being exclusively or mostly chromospheric, such as penumbral microjets (Vissers et al. 2015), shock waves (Zhao et al. 2016), spicules (Pereira et al. 2014), or jets forming in sunspot light bridges (Bharti 2015, Yang et al. 2015), are associated with significant TR emission. Conversely, many of the TR phenomena seen in SiIV 140.2-nm spectroheliograms (formed around 80,000 K under equilibrium conditions; **Figure 11**) have a direct counterpart or are driven in the chromosphere. Rapidly forming linear features shooting out from the network or plage have been called network "jets" (Tian et al. 2014) but are intimately connected to spicule formation (see below). Low-lying loops in active regions are continuously replenished from below as new flux emerges, sequentially traversing the chromosphere, TR, and corona. These highly dynamic loops are often the site of strong brightenings and bidirectional flows (as revealed by SiIV spectra), which is indicative of the magnetic reconnection that naturally occurs when large new flux systems interact with preexisting flux. Low-lying, compact loops are also seen in quiet Sun (Hansteen et al. 2014), resembling the long-hypothesized "unresolved fine-scale structure" (UFS) that had been invoked to explain the puzzling lack of correlations between TR and coronal spectral lines as observed with Skylab (Feldman 1983). From comparison with numerical simulations in which such quiet Sun loops naturally occur (Hansteen et al. 2014), it remains unclear whether they form through braiding on short loops or in response to flux emergence, as suggested by observations of very high velocities and chromospheric counterparts (Pereira et al. 2018). Footpoint regions of coronal loops (plage or network) are riddled with bright, round dots that rapidly move in the plane of the sky (Skogsrud et al. 2016): This bright TR emission is caused by slow-mode magnetoacoustic

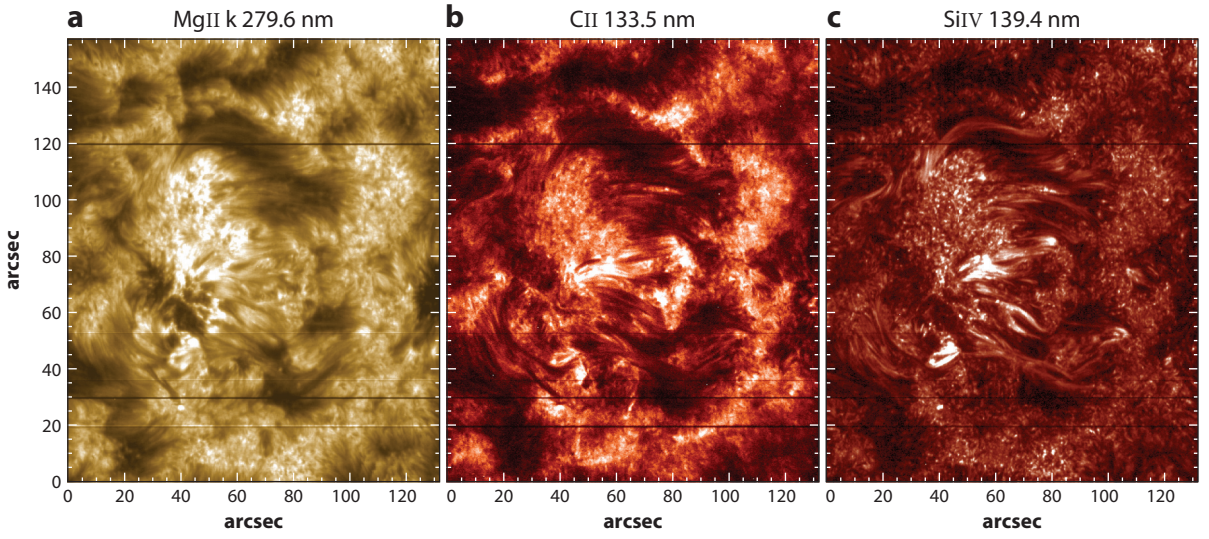

Figure 11

IRIS spectroheliograms of an active region (obtained on September 27, 2013) show the substantial morphological similarities between (*a*) the upper chromosphere, (*b*) the low transition region, and (*c*) the low middle transition region, with plage footpoints of coronal loops (*bright regions*, *mottled by round features* in SiIV, caused by magnetoacoustic shocks), disk counterparts of type II spicules (*linear features* protruding from plage), and low-lying loops associated with emerging flux (in the *middle* of the field of view). Abbreviation: IRIS, Interface Region Imaging Spectrograph.

shocks that propagate upward from the chromosphere and continuously buffet not only the TR but also the coronal loops themselves (Bryans et al. 2016, De Pontieu et al. 2017a).

These close dynamic connections imply that the evolution of lower TR plasma cannot be properly understood without accounting for the upper chromosphere, which is replete with jets, shocks, and waves, many of which propagate upward. The nature of the TR emission varies from hot features that are adjacent (e.g., on what looks like neighboring field lines) to the chromospheric phenomenon (e.g., spicules) to classical TR emission at the boundary between chromosphere and corona (shocks, jets in light bridges). Comparisons between simulations and IRIS observations show that the tight coupling between chromospheric and TR dynamics and the long ionization and recombination timescales for some elements imply that nonequilibrium ionization effects can dominate the interpretation of TR emission (e.g., Olluri et al. 2015, Martínez-Sykora et al. 2016b). Some of these results may provide a natural explanation for the long-known and puzzling correlations between TR intensity and nonthermal line broadening that were discovered with Skylab. Similarly, the nonequilibrium ionization state in the TR may affect the anomalously high SiIV/OIV intensity ratio (Martínez-Sykora et al. 2016b) because the ions are formed at different temperatures than expected from equilibrium ionization.

The impact of spicules on the outer atmosphere appears to be substantial. The slow-mode magnetoacoustic shocks that drive type I spicules lead to strong brightenings and increased nonthermal broadening (Skogsrud et al. 2016) in TR spectral lines and propagating coronal disturbances in coronal loops (Bryans et al. 2016; see also below). These shocks should be accounted for when considering, for example, the long-standing problem of pervasive redshifts and nonthermal line broadening in TR lines (Mariska 1992). The impact of type II spicules appears to be different. Statistical studies using the broad thermal coverage of IRIS, *Hinode*, and ground-based data suggest that significant heating occurs up to SiIV temperatures (80,000 K under equilibrium conditions)

in at least a subset of threads in type II spicules along their whole length as they expand into the coronal volume (Pereira et al. 2014, Skogsrud et al. 2015), providing a possible explanation for the rapid fading of spicules in some chromospheric bandpasses. An alternative idea is for decreasing densities with time, leading to loss of opacity (Beck et al. 2016), although it is unclear how this would lead to the observed increase in TR emission seen in IRIS observations. A scenario that can combine these various constraints is when spicules consist of multiple threads (Skogsrud et al. 2014) in which densities decrease with time (compatible with observations and numerical models; Beck et al. 2016, Martínez-Sykora et al. 2017b) and some threads are heated all the way to TR temperatures (De Pontieu et al. 2017a) and others stay at chromospheric temperatures (**Figure 10**). Becaues the heating to TR temperatures often occurs during the later phase of the spicule, their potential impact on the pervasive redshifts is also of interest.

The apparent heating to at least transition temperatures (SiIV, 80,000 K) along parts of the spicule provides strong constraints on theoretical models for spicule formation. Similar constraints come from the rapid propagation of the TR counterparts of spicules at 50–300 km s^{-1} (Tian et al. 2014, Narang et al. 2016), which has led to suggestions that they may provide significant mass flux into the solar wind. However, the apparent speeds in the plane of the sky are much faster than the line-of-sight flows measured from observed Doppler shifts (Rouppe van der Voort et al. 2015), suggesting that these apparent motions are not caused (fully?) by flows. Comparison with numerical simulations suggests that these apparent motions may in fact be caused by the rapid propagation (at Alfvénic speeds) and dissipation through ambipolar diffusion of electrical currents along existing chromospheric spicules (De Pontieu et al. 2017a). These currents do not cover the whole spicule but appear to shoot through some threads of the spicules, which is compatible with observational suggestions of mixed chromospheric and TR threads.

Although spicules clearly play a key role in the dynamics of the TR, their role in the corona and solar wind is not settled. Coordinated IRIS and SDO/AIA observations indicate that spicules are at the root of many coronal propagating disturbances (Pant et al. 2015, Samanta et al. 2015, Bryans et al. 2016), with the remnants of the shock waves that drive spicules thought to be responsible for the slight density perturbations associated with the coronal disturbances. However, a study by De Pontieu et al. (2017a) has provided some evidence for a connection between spicules and the formation of coronal loop strands. Comparison with the recent numerical model of spicule formation by Martínez-Sykora et al. (2017b) provides support for these observations and indicates that substantial heating of coronal plasma may be associated with spicules through the dissipation of currents, injection of hot plasma, and generation of Alfvénic waves (and their subsequent dissipation). It remains unclear how common this process is and whether it plays a dominant role in coronal heating. Simultaneous high-resolution chromospheric, TR, and coronal spectroscopy, as well as more advanced models (that can properly treat the dissipation of coronal Alfvénic waves) will be required to settle this issue.

The strong connectivity between chromosphere and corona also works in the downward direction. In particular, the lower atmosphere can be significantly impacted by coronal nanoflares, small-scale events thought to be responsible for heating in the corona in which magnetic energy is converted into kinetic, thermal, and nonthermal energy. The nonthermal particles (or electron beams) generated by nanoflares are thought to propagate downward and impact on the dense plasma in the TR and chromosphere, where they can lead to strong and short-lived brightenings (Testa et al. 2014). Recent numerical simulations suggest that the observed properties in the chromosphere and TR can provide strict constraints on the properties of the electron beams (Polito et al. 2018). The chromospheric aspects of this scenario have not yet been fully explored, on both the observational and theoretical sides, but will provide even stronger constraints on our understanding of nonthermal energy generation and deposition in the solar atmosphere.

SUMMARY POINTS

1. The chromosphere is a key region in the solar atmosphere, at the interface between the surface and its atmosphere, requiring more mechanical energy than the corona.
2. New high-resolution observatories, powerful focal instruments, adaptive optics and advanced image restoration techniques, and seeing-free measurements from space have revealed the chromosphere's highly dynamic nature on timescales of 5 s or less and fine structuring on subarcsecond scales.
3. Observations have provided deep insight into long-known chromospheric phenomena such as bright grains, spicules, Ellerman bombs, or fibrils, and discovered new features such as penumbral microjets or UV bursts.
4. Classical semiempirical 1D atmospheric models do not properly capture the complex radiative, magnetic, and plasma physics environment of the chromosphere.
5. Advances in 3D radiative MHD models have been critical for the interpretation of observations, leading to novel understanding of the physical processes that dominate the dynamics and energetics of the chromosphere:
   (a) **Magnetoacoustic slow-mode shock waves** permeate the chromosphere, both in the internetwork quiet Sun and in strong magnetic field regions like network and plage. They play a key role in the momentum and energy balance of the chromosphere, driving ionization out of equilibrium and tightly connecting the chromospheric and TR dynamics.
   (b) **Transverse magnetic waves** are ubiquitous in the chromosphere, carrying a large energy flux upward into the outer atmosphere, with their role in the local energy balance unclear despite tantalizing hints of mode conversion (e.g., through resonant absorption) and dissipation.
   (c) **Reconnection** plays a key role in the chromosphere, e.g., during large-scale flux emergence (causing, e.g., UV bursts) and likely also during heating resulting from gentle braiding of magnetic field driven by magnetoconvection.
   (d) Weak collisional coupling between ions and neutrals can lead to enhanced **ambipolar resistivity**, possibly causing substantial heating in the chromosphere, affecting flux emergence, and helping drive spicules.
   (e) **Lower transition region emission** is strongly impacted by the dominant chromospheric dynamics and heating events like jets, waves, and reconnection.
   (f) **Close coupling between corona and chromosphere** is seen at coronal loop footpoints, e.g., through nonthermal particle deposition from nanoflares.

FUTURE ISSUES

1. Major new insights into a host of unresolved issues are expected in the next decade:
   (a) the role of waves and reconnection in driving chromospheric heating and dynamics,
   (b) the impact of small-scale magnetic fields on the chromospheric energy balance,
   (c) the formation and impact of spicules,
   (d) the coupling between chromosphere, transition region, and corona, and
   (e) the mechanisms driving the first ionization potential effect and coronal heating.

2. These advances will be facilitated by a combination of the following factors:
   (a) exciting new instrumentation providing unprecedented high-resolution views of the chromosphere that include measurements of temperature (ALMA) and magnetic field (4-m-aperture DKIST telescope, and novel, integral-field-unit spectropolarimetric instruments at many telescopes),
   (b) novel inversion codes (handling nonlocal thermodynamic equilibrium, partial redistribution radiative transfer), which will dramatically expand diagnostic capabilities for the chromosphere, facilitated by drastic improvements in computational efficiency through machine learning, and
   (c) major improvements in advanced numerical modeling through inclusion of more complex physical mechanisms (e.g., increased resolution, larger domains, more realistic magnetic fields, multifluid approaches) to address important discrepancies between current models and observations that suggest a lack of small-scale motions and chromospheric heating in current models.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This research was supported by the Research Council of Norway through its Centres of Excellence scheme, project number 262622, and through grants of computing time from the Programme for Supercomputing and NASA's High End Computing division (s1061). B.D.P. was supported by NASA contract NNG09FA40C (IRIS). We thank Luc Rouppe van der Voort, David Bühler, Jayant Joshi, and Jaime de la Cruz Rodriguez for making observations from SST available, and Ada Ortiz for providing flux emergence figures. The Swedish 1-m Solar Telescope is operated on the island of La Palma by the Institute for Solar Physics (ISP) of Stockholm University in the Spanish Observatorio del Roque de los Muchachos of the Instituto de Astrofísica de Canarias. ISP is supported by a grant for research infrastructures of national importance from the Swedish Research Council (registration number 2017-00625). IRIS is a NASA small explorer mission developed and operated by the Lockheed Martin Solar and Astrophysics Laboratory with mission operations executed at NASA Ames Research Center and major contributions to downlink communications funded by the European Space Agency and the Norwegian Space Centre.

LITERATURE CITED

Acheson DJ. 1979. *Sol. Phys.* 62:23–50
Alissandrakis CE, Patsourakos S, Nindos A, Bastian TS. 2017. *Astron. Astrophys.* 605:A78
Antolin P, Okamoto TJ, De Pontieu B, et al. 2015. *Ap. J.* 809:72
Antolin P, Schmit D, Pereira TMD, et al. 2018. *Ap. J.* 856:44
Archontis V, Hansteen V. 2014. *Ap. J. Lett.* 788:L2
Asensio Ramos A, de la Cruz Rodríguez J. 2015. *Astron. Astrophys.* 577:A140
Asplund M, Grevesse N, Sauval AJ, Scott P. 2009. *Annu. Rev. Astron. Astrophys.* 47:481–522
Ballester JL, Alexeev I, Collados M, et al. 2018. *Space Sci. Rev.* 214:58
Barczynski K, Peter H, Chitta LP, Solanki SK. 2018. *Astron. Astrophys.* 619:A5

Barczynski K, Peter H, Savage SL. 2017. *Astron. Astrophys.* 599:A137
Bastian TS, Chintzoglou G, De Pontieu B, et al. 2017. *Ap. J. Lett.* 845:L19
Bastian TS, Chintzoglou G, De Pontieu B, et al. 2018. *Ap. J. Lett.* 860:L16
Beck C, Khomenko E, Rezaei R, Collados M. 2009. *Astron. Astrophys.* 507:453–67
Beck C, Rezaei R, Puschmann KG. 2012. *Astron. Astrophys.* 544:A46
Beck C, Rezaei R, Puschmann KG. 2013a. *Astron. Astrophys.* 549:A24
Beck C, Rezaei R, Puschmann KG. 2013b. *Astron. Astrophys.* 553:A73
Beck C, Rezaei R, Puschmann KG, Fabbian D. 2016. *Sol. Phys.* 291:2281–328
Beckers JM. 1968. *Sol. Phys.* 3:367–433
Beckers JM. 1972. *Annu. Rev. Astron. Astrophys.* 10:73–100
Beeck B, Collet R, Steffen M, et al. 2012. *Astron. Astrophys.* 539:A121
Bel N, Leroy B. 1977. *Astron. Astrophys.* 55:239–43
Bello González N, Franz M, Martínez Pillet V, et al. 2010. *Ap. J. Lett.* 723:L134–38
Berger TE, De Pontieu B, Schrijver CJ, Title AM. 1999. *Ap. J. Lett.* 519:L97–100
Bharti L. 2015. *MNRAS* 452:L16–20
Bjørgen JP, Sukhorukov AV, Leenaarts J, et al. 2018. *Astron. Astrophys.* 611:A62
Bogdan TJ, Carlsson M, Hansteen VH, et al. 2003. *Ap. J.* 599:626–60
Brady CS, Arber TD. 2016. *Ap. J.* 829:80
Bryans P, McIntosh SW, De Moortel I, De Pontieu B. 2016. *Ap. J. Lett.* 829:L18
Buehler D, Lagg A, Solanki SK, van Noort M. 2015. *Astron. Astrophys.* 576:A27
Carlsson M, Hansteen VH, De Pontieu B, et al. 2007. *Publ. Astron. Soc. Jpn.* 59:663–68
Carlsson M, Hansteen VH, Gudiksen BV, et al. 2016. *Astron. Astrophys.* 585:A4
Carlsson M, Judge PG, Wilhelm K. 1997. *Ap. J. Lett.* 486:L63
Carlsson M, Leenaarts J. 2012. *Astron. Astrophys.* 539:A39
Carlsson M, Leenaarts J, De Pontieu B. 2015. *Ap. J. Lett.* 809:L30
Carlsson M, Stein RF. 1992. *Ap. J. Lett.* 397:L59–62
Carlsson M, Stein RF. 1995. *Ap. J. Lett.* 440:L29–32
Carlsson M, Stein RF. 1997. *Ap. J.* 481:500–14
Carlsson M, Stein RF. 2002. *Ap. J.* 572:626–35
Cauzzi G, Reardon K, Rutten RJ, et al. 2009. *Astron. Astrophys.* 503:577–87
Centeno R. 2012. *Ap. J.* 759:72
Centeno R, Socas-Navarro H, Lites B, et al. 2007. *Ap. J. Lett.* 666:L137–40
Chatterjee P, Hansteen V, Carlsson M. 2016. *Phys. Rev. Lett.* 116:101101
Cheung MCM, Isobe H. 2014. *Living Rev. Solar Phys.* 11:3
Cranmer S, Winebarger A. 2019. *Annu. Rev. Astron. Astrophys.* 57:157–87
de la Cruz Rodríguez J, De Pontieu B, Carlsson M, Rouppe van der Voort LHM. 2013. *Ap. J. Lett.* 764:L11
de la Cruz Rodríguez J, Leenaarts J, Asensio Ramos A. 2016. *Ap. J. Lett.* 830:L30
de la Cruz Rodríguez J, Socas-Navarro H. 2011. *Astron. Astrophys.* 527:L8
de la Cruz Rodríguez J, van Noort M. 2017. *Space Sci. Rev.* 210:109–43
De Pontieu B. 2002. *Ap. J.* 569:474–86
De Pontieu B, Carlsson M, Rouppe van der Voort LHM, et al. 2012. *Ap. J. Lett.* 752:L12
De Pontieu B, De Moortel I, Martinez-Sykora J, McIntosh SW. 2017a. *Ap. J. Lett.* 845:L18
De Pontieu B, Erdélyi R, James SP. 2004. *Nature* 430:536–39
De Pontieu B, Hansteen VH, Rouppe van der Voort L, et al. 2007a. *Ap. J.* 655:624–41
De Pontieu B, Martínez-Sykora J, Chintzoglou G. 2017b. *Ap. J. Lett.* 849:L7
De Pontieu B, McIntosh S, Hansteen VH, et al. 2007b. *Publ. Astron. Soc. Jpn.* 59:655–62
De Pontieu B, McIntosh S, Martinez-Sykora J, et al. 2015. *Ap. J. Lett.* 799:L12
De Pontieu B, McIntosh SW, Carlsson M, et al. 2007c. *Science* 318:1574–77
De Pontieu B, Rouppe van der Voort L, McIntosh SW, et al. 2014a. *Science* 346:1255732
De Pontieu B, Title AM, Lemen JR, et al. 2014b. *Sol. Phys.* 289:2733–79
Dere KP, Landi E, Mason HE, et al. 1997. *Astron. Astrophys. Suppl.* 125:149–73
Fan Y. 2009. *Living Rev. Solar Phys.* 6:4

Feldman U. 1983. *Ap. J.* 275:367–73
Fletcher AC, Dimant YS, Oppenheim MM, Fontenla JM. 2018. *Ap. J.* 857:129
Fontenla JM. 2005. *Astron. Astrophys.* 442:1099–103
Fontenla JM, Avrett EH, Loeser R. 1990. *Ap. J.* 355:700–18
Fontenla JM, Avrett EH, Loeser R. 1991. *Ap. J.* 377:712–25
Fontenla JM, Avrett EH, Loeser R. 1993. *Ap. J.* 406:319–45
Fossum A, Carlsson M. 2005a. *Ap. J.* 625:556–62
Fossum A, Carlsson M. 2005b. *Nature* 435:919–21
Galsgaard K, Archontis V, Moreno-Insertis F, Hood AW. 2007. *Ap. J.* 666:516–31
Gogoberidze G, Voitenko Y, Poedts S, De Keyser J. 2014. *MNRAS* 438:3568–76
Golding TP, Carlsson M, Leenaarts J. 2014. *Ap. J.* 784:30
Golding TP, Leenaarts J, Carlsson M. 2016. *Ap. J.* 817:125
Goode PR, Coulter R, Gorceix N, et al. 2010. *Astron. Nachr.* 331:620–23
Gošić M, de la Cruz Rodríguez J, De Pontieu B, et al. 2018. *Ap. J.* 857:48
Gudiksen BV, Carlsson M, Hansteen VH, et al. 2011. *Astron. Astrophys.* 531:A154
Hansteen V, De Pontieu B, Carlsson M, et al. 2014. *Science* 346:1255757
Hansteen VH, Archontis V, Pereira TMD, et al. 2017. *Ap. J.* 839:22
Hansteen VH, Carlsson M, Gudiksen B. 2007. In *The Physics of Chromospheric Plasmas*, ed. P Heinzel, I Dorotovič, RJ Rutten, *ASP Conf. Ser.* 368:107–114. San Francisco: ASP
Hansteen VH, De Pontieu B, Rouppe van der Voort L, et al. 2006. *Ap. J. Lett.* 647:L73–76
Hasan SS, van Ballegooijen AA. 2008. *Ap. J.* 680:1542–52
Hayek W, Asplund M, Carlsson M, et al. 2010. *Astron. Astrophys.* 517:A49
Heggland L, De Pontieu B, Hansteen VH. 2007. *Ap. J.* 666:1277–83
Henriques VMJ, Kuridze D, Mathioudakis M, Keenan FP. 2016. *Ap. J.* 820:124
Heyvaerts J, Priest ER, Rust DM. 1977. *Ap. J.* 216:123–37
Hollweg JV. 1982. *Ap. J.* 257:345–53
Hotta H, Rempel M, Yokoyama T. 2015. *Ap. J.* 803:42
Iijima H, Yokoyama T. 2017. *Ap. J.* 848:38
Ionson JA. 1978. *Ap. J.* 226:650–73
Ishikawa R, Tsuneta S. 2009. *Astron. Astrophys.* 495:607–12
Ishikawa R, Tsuneta S, Ichimoto K, et al. 2008. *Astron. Astrophys.* 481:L25–28
Isobe H, Proctor MRE, Weiss NO. 2008. *Ap. J. Lett.* 679:L57
Iwai K, Loukitcheva M, Shimojo M, et al. 2017a. *Ap. J. Lett.* 841:L20
Iwai K, Shimojo M, Asayama S, et al. 2017b. *Sol. Phys.* 292:22
Jess DB, Morton RJ, Verth G, et al. 2015. *Space Sci. Rev.* 190:103–61
Kobayashi K, Cirtain J, Winebarger AR, et al. 2014. *Sol. Phys.* 289:4393–412
Kosugi T, Matsuzaki K, Sakao T, et al. 2007. *Sol. Phys.* 243:3–17
Krasnoselskikh V, Vekstein G, Hudson HS, et al. 2010. *Ap. J.* 724:1542–50
Kuridze D, Henriques V, Mathioudakis M, et al. 2015. *Ap. J.* 802:26
Kuridze D, Zaqarashvili TV, Henriques V, et al. 2016. *Ap. J.* 830:133
Landi E, Del Zanna G, Young PR, et al. 2006. *Ap. J. Suppl.* 162:261–80
Langangen Ø, Carlsson M. 2009. *Ap. J.* 696:1892–98
Leenaarts J, Carlsson M, Hansteen V, Rutten RJ. 2007. *Astron. Astrophys.* 473:625–32
Leenaarts J, Carlsson M, Rouppe van der Voort L. 2012. *Ap. J.* 749:136
Leenaarts J, de la Cruz Rodríguez J, Danilovic S, et al. 2018. *Astron. Astrophys.* 612:A28
Leenaarts J, de la Cruz Rodríguez J, Kochukhov O, Carlsson M. 2014. *Ap. J. Lett.* 784:L17
Leenaarts J, Pereira TMD, Carlsson M, et al. 2013a. *Ap. J.* 772:89
Leenaarts J, Pereira TMD, Carlsson M, et al. 2013b. *Ap. J.* 772:90
Leenaarts J, Wedemeyer-Böhm S. 2006. *Astron. Astrophys.* 460:301–7
Lemen JR, Title AM, Akin DJ, et al. 2012. *Sol. Phys.* 275:17–40
Lin HH, Carlsson M. 2015. *Ap. J.* 813:34
Lin HH, Carlsson M, Leenaarts J. 2017. *Ap. J.* 846:40

Lites BW. 2009. *Space Sci. Rev.* 144:197–212
Loukitcheva M, Solanki SK, Carlsson M, White SM. 2015. *Astron. Astrophys.* 575:A15
Loukitcheva M, White SM, Solanki SK, et al. 2017a. *Astron. Astrophys.* 601:A43
Loukitcheva MA, Iwai K, Solanki SK, et al. 2017b. *Ap. J.* 850:35
Maltby P, Avrett EH, Carlsson M, et al. 1986. *Ap. J.* 306:284–303
Mariska JT. 1992. *The Solar Transition Region*. Cambridge, UK: Cambridge Univ. Press
Martínez González MJ, Bellot Rubio LR. 2009. *Ap. J.* 700:1391–403
Martínez González MJ, Collados M, Ruiz Cobo B, Solanki SK. 2007. *Astron. Astrophys.* 469:L39–42
Martinez Pillet V, Garcia Lopez RJ, del Toro Iniesta JC, et al. 1990. *Ap. J. Lett.* 361:L81–85
Martínez-Sykora J, De Pontieu B, Carlsson M, Hansteen V. 2016a. *Ap. J. Lett.* 831:L1
Martínez-Sykora J, De Pontieu B, Carlsson M, et al. 2017a. *Ap. J.* 847:36
Martínez-Sykora J, De Pontieu B, Hansteen V. 2012. *Ap. J.* 753:161
Martínez-Sykora J, De Pontieu B, Hansteen V, Carlsson M. 2015a. *Philos. Trans. R. Soc. Lond. Ser. A* 373:40268
Martínez-Sykora J, De Pontieu B, Hansteen VH, Gudiksen B. 2016b. *Ap. J.* 817:46
Martínez-Sykora J, De Pontieu B, Hansteen VH, et al. 2017b. *Science* 356:1269–72
Martínez-Sykora J, De Pontieu B, Leenaarts J, et al. 2013. *Ap. J.* 771:66
Martínez-Sykora J, Hansteen V, DePontieu B, Carlsson M. 2009. *Ap. J.* 701:1569–81
Martínez-Sykora J, Hansteen V, Moreno-Insertis F. 2011. *Ap. J.* 736:9
Martínez-Sykora J, Rouppe van der Voort L, Carlsson M, et al. 2015b. *Ap. J.* 803:44
Michalitsanos AG. 1973. *Sol. Phys.* 30:47–61
Mihalas D. 1978. *Stellar Atmospheres*. San Francisco: W.H. Freeman and Company. 2nd ed.
Moreno-Insertis F. 2007. In *New Solar Physics with Solar-B Mission*, ed. K Shibata, S Nagata, T Sakurai. *ASP Conf. Ser.* 369:335–46. San Francisco: ASP
Narang N, Arbacher RT, Tian H, et al. 2016. *Sol. Phys.* 291:1129–42
Nordlund A. 1982. *Astron. Astrophys.* 107:1–10
Okamoto TJ, Antolin P, De Pontieu B, et al. 2015. *Ap. J.* 809:71
Okamoto TJ, De Pontieu B. 2011. *Ap. J. Lett.* 736:L24
Olluri K, Gudiksen BV, Hansteen VH, De Pontieu B. 2015. *Ap. J.* 802:5
Ortiz A, Bellot Rubio LR, Hansteen VH, et al. 2014. *Ap. J.* 781:126
Pant V, Dolla L, Mazumder R, et al. 2015. *Ap. J.* 807:71
Pariat E, Aulanier G, Schmieder B, et al. 2004. *Ap. J.* 614:1099–112
Parker EN. 1988. *Ap. J.* 330:474–79
Pereira TMD, Asplund M, Collet R, et al. 2013a. *Astron. Astrophys.* 554:A118
Pereira TMD, Carlsson M, De Pontieu B, Hansteen V. 2015. *Ap. J.* 806:14
Pereira TMD, De Pontieu B, Carlsson M. 2012. *Ap. J.* 759:18
Pereira TMD, De Pontieu B, Carlsson M, et al. 2014. *Ap. J. Lett.* 792:L15
Pereira TMD, Leenaarts J, De Pontieu B, et al. 2013b. *Ap. J.* 778:143
Pereira TMD, Rouppe van der Voort L, Hansteen VH, De Pontieu B. 2018. *Astron. Astrophys.* 611:L6
Polito V, Testa P, Allred J, et al. 2018. *Ap. J.* 856:178
Quintero Noda C, Uitenbroek H, Katsukawa Y, et al. 2017. *MNRAS* 470:1453–61
Rathore B, Carlsson M. 2015. *Ap. J.* 811:80
Rathore B, Carlsson M, Leenaarts J, De Pontieu B. 2015a. *Ap. J.* 811:81
Rathore B, Pereira TMD, Carlsson M, De Pontieu B. 2015b. *Ap. J.* 814:70
Reale F. 2014. *Living Rev. Solar Phys.* 11:4
Rempel M. 2014. *Ap. J.* 789:132
Rezaei R, Schlichenmaier R, Beck CAR, et al. 2007. *Astron. Astrophys.* 466:1131–44
Rouppe van der Voort L, De Pontieu B, Pereira TMD, et al. 2015. *Ap. J. Lett.* 799:L3
Rouppe van der Voort L, Leenaarts J, De Pontieu B, et al. 2009. *Ap. J.* 705:272–84
Rouppe van der Voort LHM, De Pontieu B, Hansteen VH, et al. 2007. *Ap. J. Lett.* 660:L169–72
Rutten RJ. 2002. *J. Astron. Data* 8:23–54
Samanta T, Pant V, Banerjee D. 2015. *Ap. J. Lett.* 815:L16

Scharmer GB, Bjelksjo K, Korhonen TK, et al. 2003. In *Innovative Telescopes and Instrumentation for Solar Astrophysics*, ed. SL Keil, SV Avakyan. *SPIE Conf. Ser.* 4853. Bellingham, WA: SPIE. **https://doi.org/10.1117/12.460377**
Scharmer GB, Narayan G, Hillberg T, et al. 2008. *Ap. J. Lett.* 689:L69–72
Schmidt W, von der Lühe O, Volkmer R, et al. 2012. *Astron. Nachr.* 333:796
Schmit D, Bryans P, De Pontieu B, et al. 2015. *Ap. J.* 811:127
Sekse DH, Rouppe van der Voort L, De Pontieu B. 2012. *Ap. J.* 752:108
Sekse DH, Rouppe van der Voort L, De Pontieu B. 2013a. *Ap. J.* 764:164
Sekse DH, Rouppe van der Voort L, De Pontieu B, Scullion E. 2013b. *Ap. J.* 769:44
Shimojo M, Bastian TS, Hales AS, et al. 2017a. *Sol. Phys.* 292:87
Shimojo M, Hudson HS, White SM, et al. 2017b. *Ap. J. Lett.* 841:L5
Skartlien R. 2000. *Ap. J.* 536:465–80
Skogsrud H, Rouppe van der Voort L, De Pontieu B. 2014. *Ap. J. Lett.* 795:L23
Skogsrud H, Rouppe van der Voort L, De Pontieu B. 2016. *Ap. J.* 817:124
Skogsrud H, Rouppe van der Voort L, De Pontieu B, Pereira TMD. 2015. *Ap. J.* 806:170
Sobotka M, Švanda M, Jurčak J, et al. 2014. *Cent. Eur. Astrophys. Bull.* 38:53–58
Solanki SK, Barthol P, Danilovic S, et al. 2010. *Ap. J. Lett.* 723:L127–33
Solanki SK, Riethmüller TL, Barthol P, et al. 2017. *Ap. J. Suppl.* 229:2
Sollum E. 1999. *Hydrogen ionization in the solar atmosphere: exact and simplified treatments*. MSc Thesis, Inst. Theor. Astrophys., Univ. Oslo, Nor.
Steffens S, Deubner FL, Fleck B, Wilhelm K. 1997. In *Proceedings of the 5th SOHO Workshop: The Corona and Solar Wind Near Minimum Activity*, ed. A Wilson. *ESA Spec. Pub.* 404:685. Noordwijk, Neth.: ESA
Sterling AC. 2000. *Sol. Phys.* 196:79–111
Sterling AC, Moore RL. 2016. *Ap. J. Lett.* 828:L9
Suematsu Y. 1990. In *Progress of Seismology of the Sun and Stars*, *Lect. Notes Phys.*, Vol. 367, ed. Y Osaki, H Shibahashi, pp. 211–14. Berlin: Springer-Verlag
Suematsu Y, Wang H, Zirin H. 1995. *Ap. J.* 450:411
Testa P, De Pontieu B, Allred J, et al. 2014. *Science* 346:1255724
Tian H, DeLuca EE, Cranmer SR, et al. 2014. *Science* 346:1255711
Tortosa-Andreu A, Moreno-Insertis F. 2009. *Astron. Astrophys.* 507:949–67
Tsuneta S, Ichimoto K, Katsukawa Y, et al. 2008. *Sol. Phys.* 249:167–96
Ugai M, Tsuda T. 1977. *J. Plasma Phys.* 17:337–56
Uitenbroek H. 2001. *Ap. J.* 557:389–98
van Ballegooijen AA, Asgari-Targhi M, Cranmer SR, DeLuca EE. 2011. *Ap. J.* 736:3
Vernazza JE, Avrett EH, Loeser R. 1973. *Ap. J.* 184:605–32
Vernazza JE, Avrett EH, Loeser R. 1976. *Ap. J. Suppl.* 30:1–60
Vernazza JE, Avrett EH, Loeser R. 1981. *Ap. J. Suppl.* 45:635–725
Vissers GJM, Rouppe van der Voort LHM, Carlsson M. 2015. *Ap. J. Lett.* 811:L33
Wang Y, Su Y, Shen J, et al. 2018. *Ap. J.* 859:148
Wang YM. 2016. *Ap. J. Lett.* 820:L13
Wedemeyer S, Bastian T, Brajša R, et al. 2016. *Space Sci. Rev.* 200:1–73
Welsch BT. 2015. *Publ. Astron. Soc. Jpn.* 67:18
White SM, Iwai K, Phillips NM, et al. 2017. *Sol. Phys.* 292:88
Wikstøl Ø, Hansteen VH, Carlsson M, Judge PG. 2000. *Ap. J.* 531:1150–60
Wootten A, Thompson AR. 2009. *Proc. IEEE* 97:1463–71
Yang S, Zhang J, Jiang F, Xiang Y. 2015. *Ap. J. Lett.* 804:L27
Yokoyama T, Shimojo M, Okamoto TJ, Iijima H. 2018. *Ap. J.* 863:96
Yurchyshyn VB, Goode PR, Abramenko VI, et al. 2010. *Ap. J.* 722:1970–76
Zhang YZ, Shibata K, Wang JX, et al. 2012. *Ap. J.* 750:16
Zhao J, Felipe T, Chen R, Khomenko E. 2016. *Ap. J. Lett.* 830:L17
Zwaan C. 1987. *Annu. Rev. Astron. Astrophys.* 25:83–111

Annual Review of Astronomy and Astrophysics

Volume 57, 2019

Contents

Indexes

Errata

An online log of corrections to *Annual Review of Astronomy and Astrophysics* articles may be found at http://www.annualreviews.org/errata/astro