The Astrophysical Journal, 659:1673–1681, 2007 April 20

THE CORONAL HEATING PARADOX

Markus J. Aschwanden
Lockheed Martin Advanced Technology Center, Solar and Astrophysics Laboratory, Org. ADBS, Palo Alto, CA 94304; aschwanden@lmsal.com

Amy Winebarger
AAMU Physics Department, Normal, AL 35762; winebarger@physics.aamu.edu

David Tsiklauri
Institute for Materials Research, University of Salford, Greater Manchester M5 4WT, England; d.tsiklauri@salford.ac.uk

and

Hardi Peter
Kiepenheuer-Institut fur Sonnenphysik, 79104 Freiburg, Germany; peter@kis.uni-freiburg.de
Received 2006 October 4; accepted 2006 December 22

ABSTRACT

The "coronal heating problem" has been with us over 60 years, and hundreds of theoretical models have been proposed without an obvious solution in sight. In this paper we point out that observations show no evidence for local heating in the solar corona, but rather for heating below the corona in the transition region and upper chromosphere, with subsequent chromospheric evaporation as known in flares. New observational evidence for this scenario comes from (1) the temperature evolution of coronal loops, (2) the overdensity of hot coronal loops, (3) upflows in coronal loops, (4) the Doppler shift in coronal loops, (5) upward propagating waves, (6) the energy balance in coronal loops, (7) the magnetic complexity in the transition region, (8) the altitude of nanoflares and microflares, (9) the cross section of elementary loops, as well as (10) 3D MHD simulations of coronal heating. The phrase "coronal heating problem" is therefore a paradoxical misnomer for what should rather be addressed as the "chromospheric heating problem" and "coronal loop filling process." This paradigm shift substantially reduces the number of relevant theoretical models for coronal heating in active regions and the quiet Sun, but our arguments do not apply to coronal holes and the extended heliospheric corona.

Subject headings: Sun: chromosphere — Sun: corona — hydrodynamics

1. INTRODUCTION

If we wanted to heat a house in a cold climate zone, we would install a conventional wood-burning oven, a gas or oil central heating system, or solar cells on the roof. What all these heating systems have in common is a centralized heating source that redistributes the primarily generated heat throughout the house by air ducts or water pipelines. We would like to argue here that the hot temperature of the solar corona is redistributed in an analogous way, generated by a primary heating process located in the solar transition region or upper chromosphere. The point is that the heat generation does not occur in the corona per se, but rather in the chromosphere, reducing the so-called coronal heating problem to a chromospheric affair with subsequent filling of the corona. The house-heating analogy could be stretched by the argument that a warmer temperature in a house does not necessarily require a centralized heating source in the basement, but could also be obtained uniformly throughout the house by external Sun exposure. There is indeed a solar analog, in the sense that the solar corona could be uniformly heated by external waves (of subphotospheric origin), but we will see that such a scenario cannot explain many observables. If we accept that the primary heating is generated in the chromosphere/transition region, then the phrase of a "coronal heating process" becomes a paradoxical misnomer, because the hot temperature of the corona is merely conveyed by upflows of heated plasma, which fills coronal magnetic loops like conduits, ducts, or pipelines. The heating problem should then rather be addressed as the "chromosphere/transition region heating process" with subsequent "coronal loop filling" by upflows.

2. CORONAL HEATING REQUIREMENT

To localize the sources of predominant coronal heating it is most illustrative to examine a soft X-ray image of the Sun, where the total emission measure of coronal temperatures $T \gtrsim 1.0$ MK is diagnosed over the entire hot temperature range. Such a full-Sun composite image that captures the extended corona out to $r \gtrsim 2\ R_\odot$ is shown in Figure 1, obtained from *Yohkoh* observations on 1992 August 26 (Aschwanden & Acton 2001). In this image, the solar corona above the limb has been subdivided into 36 azimuthal sectors with 10° width each, and the average base density n_0, base temperature T_0, and scale heights λ_T have been measured from forward-fitting of line-of-sight integrated differential emission measure distributions (Table 1 in Aschwanden & Acton 2001). Based on these values, we estimate the heating power requirement P_H (per unit area and time) from the energy balance of the heating rate E_H (per volume and time) integrated over a vertical scale height λ_T,

$$P_H = E_H \lambda_T \approx -E_R \lambda_T = n_0^2 \Lambda(T_0) \lambda_T, \qquad (1)$$

where E_R is the radiative loss rate and the radiative loss function can be approximated by a constant $\Lambda(T) \approx 10^{-22}$ (ergs cm^3 s^{-1}) in the temperature range of $T \approx 0.5$–3 MK (Rosner et al. 1978), which covers most of the coronal temperature range. The result for the 36 sectors is shown in Figure 1 (*bottom panel*) as function of the azimuthal position. We mark those sectors with active regions with dark gray, quiet-Sun regions with light gray, and coronal holes with white in the histogram shown in Figure 1

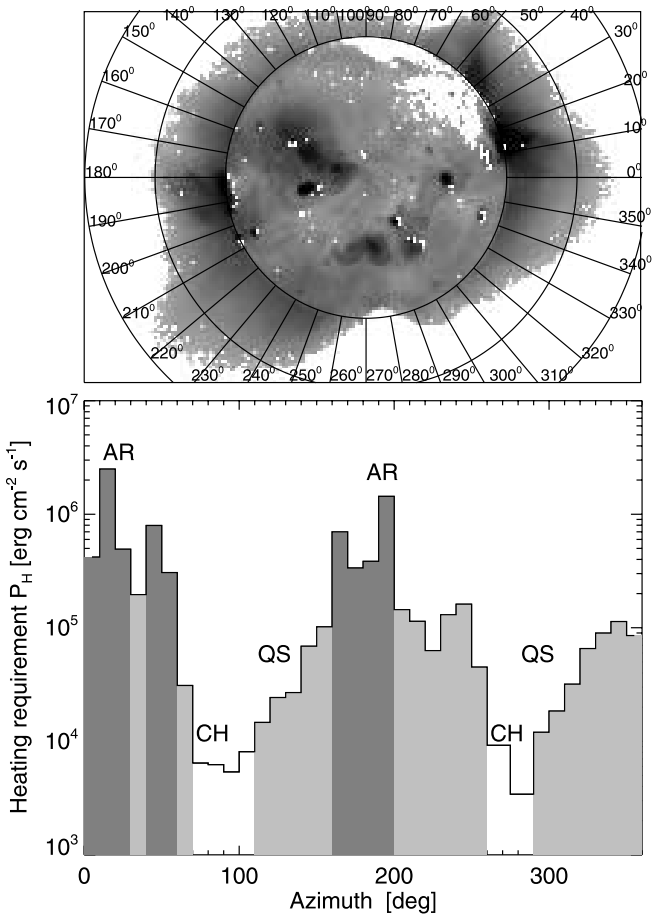

FIG. 1.—Composite soft X-ray image of the Sun observed on 1992 August 26 with *Yohkoh* (*top panel*). The histogram shows the heating rate requirement (*bottom panel*) in the 36 azimuthal sectors around the Sun. The labels indicate the locations of active regions (AR; *dark gray*), quiet-Sun regions (QS; *light gray*), and coronal holes (CH; *white*) (Aschwanden 2001).

(*bottom panel*). The heating requirement is about $2 \times 10^5 \lesssim P_H \lesssim 2 \times 10^6$ ergs cm^{-2} s^{-1} in active regions, about $1 \times 10^4 \lesssim P_H \lesssim 2 \times 10^5$ ergs cm^{-2} s^{-1} in quiet-Sun regions, and $5 \times 10^3 \lesssim P_H \lesssim 1 \times 10^4$ ergs cm^{-2} s^{-1} in coronal holes, which is in agreement with the estimates of Withbroe & Noyes (1977). If we sum the heating energy requirement in those three categories from the 36 sectors, we find that the active regions demand 82.4% of the heating requirement, the quiet-Sun regions 17.2%, and coronal holes merely 0.4%. Therefore, the total energy budget of the coronal heating problem is dominated by the heating requirement of active regions, as pointed out in Aschwanden (2001), and thus we mostly focus our arguments in § 3 on active regions loops rather than on coronal holes. Besides the radiative loss, the solar wind flux in coronal holes is the next largest energy loss component ($F_W \lesssim 7 \times 10^5$ ergs cm^{-2} s^{-1}), but is typically smaller than the radiative loss rate in active regions.

3. TEN ARGUMENTS FOR HEATING IN THE CHROMOSPHERE/TRANSITION REGION

3.1. *Temperature Evolution of Coronal Loops*

Most of the heating is required in active regions and quiet-Sun regions, which consist mostly of closed magnetic field structures. Let us compare the two scenarios of coronal heating versus chromospheric heating, as illustrated in Figure 2. If the magnetic loops are directly heated in the solar corona, the temperature of a coronal loop should increase over some time interval, as long as the heating rate exceeds the conductive and radiative cooling rate. The heating can also occur intermittently and may not be resolved in time with an observed cadence, but the average excess heating rate can produce a systematic temperature increase of the loop. A loop that is locally heated in the corona is then expected to brighten up and dim in an EUV filter first (say at $T \approx 1$–2 MK), and subsequently in soft X-ray filters (say at $T \approx 2$–4 MK), and vice versa during the cooling phase (Fig. 2, *top right panel*). However, such an evolution that includes the heating phase in EUV has never been observed (or a least not been reported in literature), which means that the heating phase is either very rapid (less than the typical observing cadence of about 1 minute) or the heating does not occur in the corona.

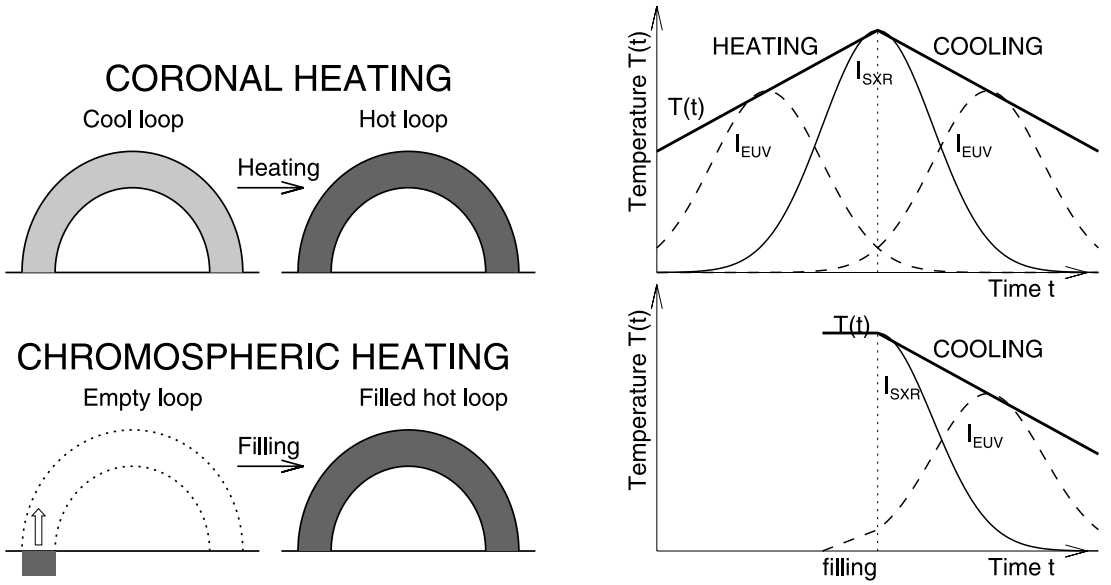

FIG. 2.—Schematic diagram of coronal heating scenario (*top panels*) vs. "chromospheric heating scenario" (*bottom panels*). The coronal heating scenario involves the heating of a preexisting overdense cool loop in the corona, while the chromospheric heating scenario fills an empty loop structure in the corona. The heating phase of the coronal heating scenario should be observable in EUV (*top right panel*), which is missing in the chromospheric filling scenario (*bottom right panel*).

On the other hand, if we consider a chromospheric evaporation scenario, where heating occurs in the upper chromosphere or lower transition region, with subsequent upflows and filling of coronal loops, we would not detect any heating phase in the coronal part of a loop. The process of chromospheric evaporation has been extensively modeled with hydrodynamic simulations for solar flares (e.g., MacNeice et al. 1984; Nagai & Emslie 1984; Fisher et al. 1984, 1985a, 1985b; Mariska & Poland 1985; Yokoyama & Shibata 1998, 2001; Hori et al. 1997, 1998; Tsiklauri et al. 2004), as well as for quiescent heating of coronal loops (Warren et al. 2002, 2003). The driver of the chromospheric evaporation process in flares can be either downward moving thermal conduction fronts or precipitating nonthermal particles, and the latter process could even work for nanoflare heating (Brown et al. 2000). For nonflare conditions, the "tectonic" photospheric random motion driven by subphotospheric convection is likely to cause small-scale reconnection processes in the transition region (e.g., Priest et al. 2002), which heat up the plasma at the footpoints of coronal loops and fill them up subsequently. Recent hydrodynamic simulations by Warren et al. (2002, 2003) employ an intermittent heating process localized at the loop footpoints, which produces a sequence of repeated heating pulses that are able to mimic the observed temperature evolution $T(s, t)$ of coronal loops. As soon as the overpressure starts to fill up the loop with hot plasma, the loop is expected to brighten rapidly in the soft X-ray filter first (during a filling-time interval that is in the order of a shock loop-crossing time, i.e., within minutes for typical coronal loops), without any detectable signal in the cooler EUV channels. Consequently, we would observe only the cooling phase, when the loop cools down from the soft X-ray filters through EUV filters, as sketched in Fig. 2 (*bottom right panel*). Such loop cooling phases from *Yohkoh* soft X-ray filters down to *TRACE* EUV filters have indeed been observed in detail for 11 cases, over time intervals of several hours, without any noticeable EUV signature of an initial heating phase (Winebarger & Warren 2005; Ugarte-Urra et al. 2006). The footpoint of a loop has been monitored in EUV during the initial rise in soft X-rays in detail in a single case (Fig. 6 in Ugarte-Urra et al. 2006), but the full loop was not detected in EUV ($T = 1.0$–1.5 MK) until it cooled down from the hotter ($T = 3$–5 MK) soft X-ray temperatures. Caveats of the "chromospheric evaporation scenario" for heating of quiescent loops are (1) the nondetection of the initial heating agent (small-scale reconnections events in the transition region) and (2) the lack of observed blueshifted upflows, which are probably to blame on insufficient time cadence, spatial, and spectral resolution.

3.2. *Overdensity of Coronal Loops*

Coronal loops are only detectable if they have a density contrast with respect to the ambient background corona. Since their brightness in EUV and soft X-rays scales linearly with the emission measure, which is proportional to the squared density, they are overdense with respect to the background corona. Overdense loops, however, have an increased radiative loss rate, and thus need to be heated stronger than the ambient low-density corona. In other words, heated loops have an overdensity with respect to the background corona, a fact that needs to be explained by every heating mechanism. However, the electron density in a coronal loop is not expected to significantly increase by a local heating process in the corona, e.g., by wave heating or magnetic reconnection (for a discussion of theoretical models, see Aschwanden 2001). The electron density in coronal loops can only be significantly enhanced by upflows of additional plasma from the loop footpoints, which requires a secondary heating process in the upper chromosphere or lower transition region. We argue that every primary heating mechanism in the corona is not able to explain the observed overdensity or emission measure excess in hot coronal loops, unless chromospheric evaporation occurs (Fig. 2, *bottom*).

The higher density of hot (soft X-ray-emitting) loops relative to the cooler (EUV-emitting) has also been statistically quantified in terms of a scaling law, i.e., $n_e(T_e) \propto T_e^2$ (Aschwanden 1999), inferred from a compilation of loop data in the temperature range of $T \approx 1$–20 MK (Pallavicini et al. 1977; Kano & Tsuneta 1995; Shimizu 1997; Aschwanden & Benz 1997; Aschwanden et al. 2000b). This observed correlation can also be explained in terms of the RTV scaling law, $n_e \propto T^2/L$ (Rosner et al. 1978; Aschwanden et al. 2006).

3.3. *Upflows in Coronal Loops*

The best direct evidence for chromospheric heating with subsequent coronal filling is observations of hot upflows from the chromosphere. Such upflows have been inferred from feature tracking in *TRACE* Fe IX/Fe X (171 Å) images in AR 8395 on 1998 December 1 (Winebarger et al. 2001). A bundle of fan-shaped loops (Fig. 3, *top left panel*) shows an increase in intensity that propagates along loop 3, which is interpreted as a tracer of apparent flows. A time sequence of four brightness profiles is shown in Figure 3 (*bottom left panel*), and a sequence of 10 brightness maps of stripes extracted along loop 3 is shown in Figure 3 (*right panel*). The projected velocity of the leading edge was measured in the range of $v \approx 5$–15 km s^{-1} for four different upflow events, each one lasting between 2 and 5 minutes. Because of projection effects, these measured velocities represent only lower limits to the actual flow speeds. These dynamic features in plasmas with a temperature of $T \approx 1.0$ MK represent upflows of heated plasma from the chromosphere into the corona, because they could not be reproduced with quasi-static changes in hydrostatic loop models without flows (Winebarger et al. 2001). Also the hydrodynamic modeling of three other coronal loops required unidirectional flow speeds in the order of $v \approx 10$–40 km s^{-1} (Petrie et al. 2003). Direct detection of (supersonic) upflow speeds with Doppler shift measurements, however, is notoriously difficult because of insufficient spatial resolution to separate the superimposed upflows and downflows (with one laudable exception, see Czaykowska et al. 1999).

3.4. *Hot Blueshifts and Cool Redshifts*

Doppler shift measurements of chromospheric UV emission lines and coronal EUV emission lines (Fig. 4) revealed that hot ($T \approx 0.5$–1.5 MK) coronal lines (Ne VIII, Mg X, Fe XII) exhibit a center-to-limb behavior that is consistent with disk-center blueshifts, while the cooler ($\approx$0.05–0.5 MK) chromospheric lines (C II, C IV, Si IV, S VI, N V, O IV, O V, O VI) exhibit redshifts (Peter & Judge 1999). This systematic trend of hot upflows and cooler downflows can straightforwardly be interpreted by a heating mechanism located in the chromosphere/transition region, which fills coronal loops with hot ($T > 0.5$ MK) plasma, where it cools down, until the thermal instability sets in (below the peak temperature of the radiative loss function around $T \lesssim 0.5$ MK) and subsequently leads to catastrophic cooling with downflows (Schrijver 2001; Müller et al. 2003, 2004, 2005). Such a highly time-variable scenario would also explain the large scatter of the line shifts in time and space, which is observed to largely exceed the average values as plotted in Figure 4 (Peter 1999). Spectra synthesized from three-dimensional (3D) box simulations of the corona (§ 3.10), with the heating concentrated at

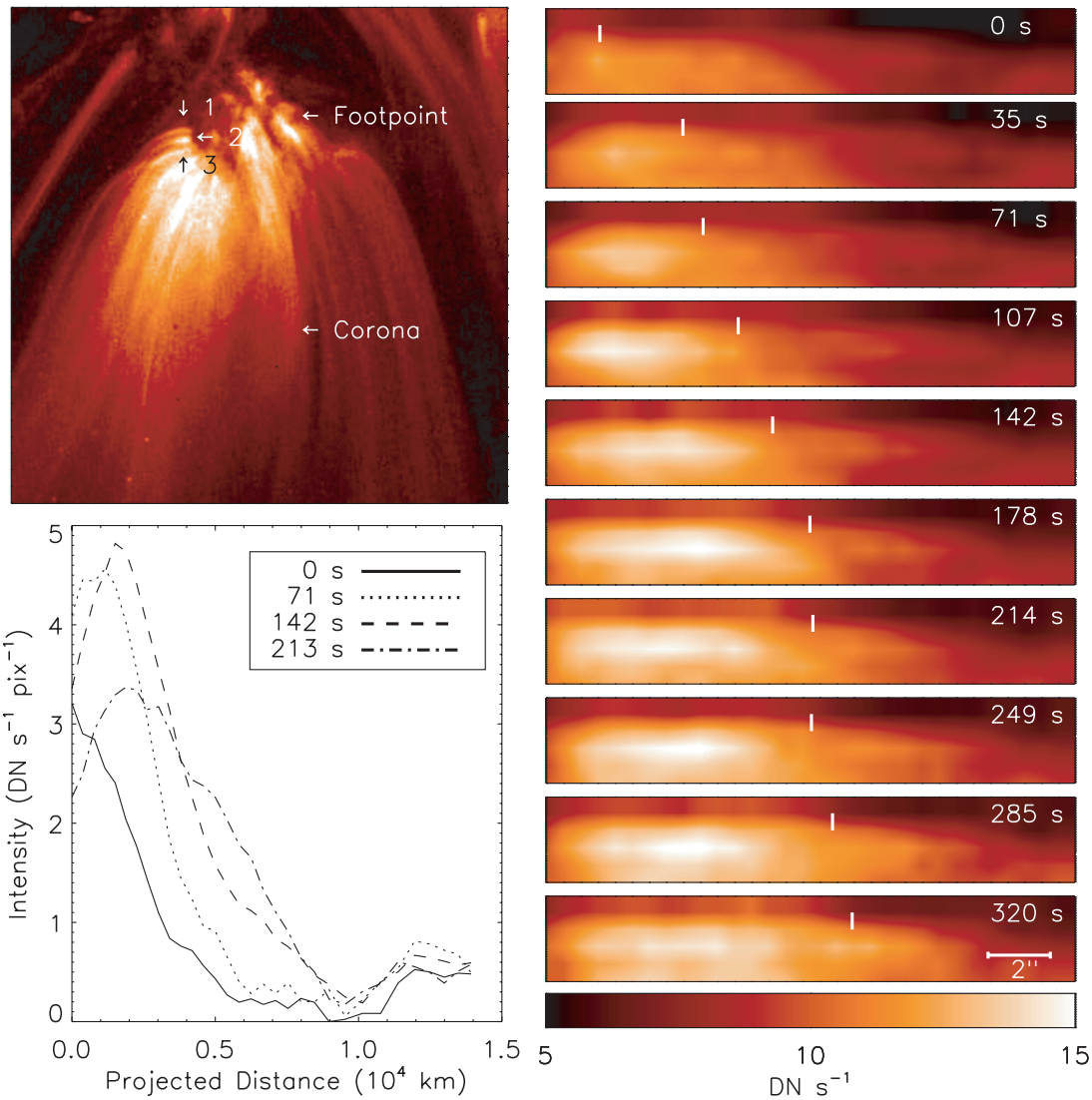

FIG. 3.—*Top left: TRACE* 171 Å image of fanlike loops in AR 8396 observed near the Sun center on 1998 December 1, 01:00:14 UT, shown with a field of view of 270 Mm. *Bottom left:* Intensity profiles along loop 3 over a projected distance of 15 Mm taken at four different times. *Right side:* Gray-scale representation of the flux along loop 3 (over a segment of 15 Mm) observed at 10 time intervals after 01:40:35 UT, with the footpoint of the fanlike loop on the left side (Winebarger et al. 2001).

low heights, reproduce the average values and the scatter of the observed redshifts in the middle transition region quite well (Peter et al. 2004, 2006). These models provide the best overall match to the transition region redshifts, so far, thus supporting coronal heating to be concentrated deep down. Thus, this cycle of hot upflows and cool downflows clearly points to the transition region as primary heating source.

3.5. *Upward Propagating Waves*

Any impulsive heating mechanism is expected to produce a disturbance of the local density and/or temperature profile, which subsequently will be smoothed out until a new pressure equilibrium is reached. The disturbance propagates with acoustic speed, or with supersonic shock speed in case of large impulsive energy depositions. Such propagating acoustic waves with speeds of $v \approx 150$ km s^{-1} (at $T \approx 1.0$ MK) have indeed been observed in diverging fanlike coronal structures in *TRACE* 171 Å data (DeMoortel et al. 2000, 2002a, 2002b, 2002c), which approximately extend over the lowest density scale height of large-scale coronal loops or open-ended flux tubes. The direction of these propagating waves has been identified to be always from the footpoints upward, while no downward propagating waves were ever detected (DeMoortel et al. 2000, 2002a, 2002b, 2002c). This preference for upward propagating disturbances strongly indicates that the associated impulsive energy depositions in coronal loops occurs near the (chromospheric) footpoints, rather than near the loop apexes. The dissipation of the upward propagating strongly damped wave trains in the fundamental harmonic mode was found to be insufficient to explain the heating of the coronal loops (DeMoortel et al. 2002c). However, as shown by Tsiklauri & Nakariakov (2001), the wide-band spectrum slows magnetoacoustic waves, consistent with currently available observations in the low-frequency part of the spectrum, can provide a rate of heat deposition sufficient to heat the loop. In this scenario, the heat would be deposited near the loop footpoints, which agrees with the current observational data.

3.6. *Energy Balance and Heating Function*

The energy balance in coronal loops has been calculated for different heating functions, such as for footpoint heating, loop-top heating, or uniform heating. The resulting temperature $T_e(s)$ and density profiles $n_e(s)$ in hydrostatic equilibrium have been

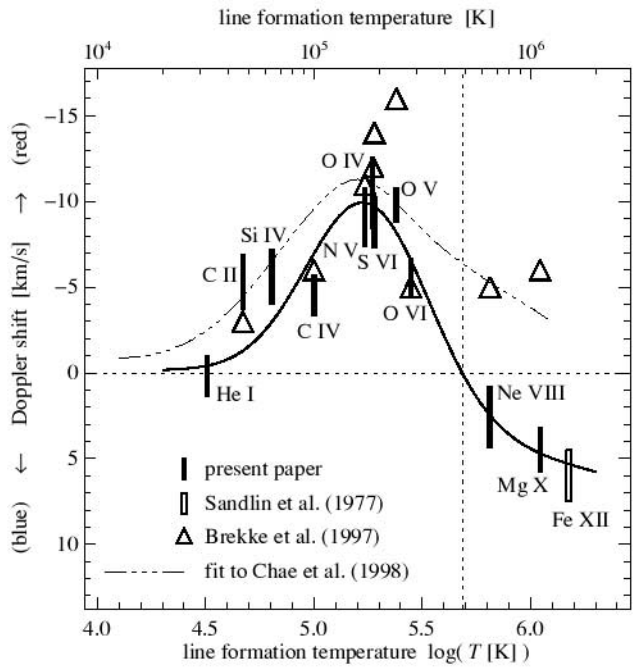

FIG. 4.—Variation of the average Doppler shift at disk center with formation temperature of the line. Cool lines with formation temperatures of $T \lesssim 0.5$ MK show redshifts, while hotter lines show blue shifts (Peter & Judge 1999).

originally calculated for a uniform (volumetric) heating rate $dE_H/dV\,dt$ by Rosner et al. (1978), for footpoint heating by Serio et al. (1981), and expressed with analytical approximations by Aschwanden & Schrijver (2002). The resulting temperature profile shows a steady temperature increase toward the loop top for uniform heating, but is essentially flat for footpoint heating (except for a very steep increase in the transition region). Fitting these equilibrium solutions to observed temperatures of coronal loops in *TRACE* data clearly demonstrated consistency with footpoint heating, with an average heating scale height of $s_H = 17 \pm 6$ Mm (Aschwanden et al. 2000a). This result can simply be interpreted in terms of dominant radiative loss, which requires predominant heating within an altitude of a half-density scale height ($\lambda_R = \lambda_T/2 \approx 23$ Mm for $T_e \approx 1.0$ MK), since the radiative loss drops off with height as the squared density [$E_R(h) \approx n_e(h)^2$; see eq. (1)]. A coronal heating function $E_H(h)$ that exactly matches the radiative loss scale height λ_H would represent an extremely restricting condition for any heating mechanism, while steady upflow of heated plasma represents a much more natural process to satisfy the energy balance. A subsonic flow crosses a heating scale height (i.e., a half-density scale height) within a time interval of $\Delta t \approx \lambda_H/v \approx 12$ minutes for a typical coronal loop ($v = 0.2c_s = 30$ km s^{-1}, $T = 1.0$ MK, and $\lambda_R = \lambda_T/2 = 22{,}000$ km), which is much shorter than the radiative cooling time,

$$\tau_R = \frac{\epsilon_{\rm th}}{dE_R/dt} = \frac{3n_e k_B T_t}{n_e^2 \Lambda(T_0)}, \qquad (2)$$

which amounts to $\tau_R \approx 1$ hr for a loop with $T_e \approx 1.0$ MK and $n_e \approx 10^9$ cm^{-3}, and progressively longer for lower densities. Therefore, steady upflows can keep radiatively cooling loops at nearly constant temperatures (in time) as observed, while a coronal heating source could not explain unidirectional flows and time-constant temperature profiles in a natural way.

3.7. *Magnetic Complexity in Transition Region*

Magnetic reconnection, driven by magnetic shear, convection, or eruption, is considered as a prime energy release mechanism for many coronal heating scenarios. For instance, magnetic reconnection in tangential discontinuities between braided and tangled coronal loops was envisioned by Parker (1988) as a viable coronal (nanoflare) heating scenario. Tangential discontinuities result from misalignment of neighbored magnetic field lines that are stirred by random footpoint motion, ultimately driven by subphotospheric convection. So, the magnetic complexity in terms of misaligned field lines is a likely measure of local reconnection events. The magnetic complexity in the transition region and upper chromosphere is, however, much greater than in the corona, because the subphotospheric convection flow transports footpoints of coronal field lines to photospheric network lanes, which produces overarching canopies in the transition region. The magnetic connectivity from coronal loops down to the photosphere is very difficult to track (Jendersie & Peter 2006), but it appears that coronal loops do not connect directly on the straightest way to the surface (DePontieu et al. 2003), but rather follow along domelike and canopy-like structures as shown in Figure 5. Such magnetic topologies have been theoretically modeled in terms of 3D null points, which naturally form in the intersection of a "spine field line" above an isolated magnetic polarity with a domelike "fan separatrix surface" formed by a surrounding circle of opposite magnetic polarity (Brown & Priest 2001). At the footpoints of coronal loops, such domelike separatrix surfaces are thought to form over the size of photospheric granulation cells, supergranulation cells, and network structures, which have typical sizes of $l \approx 1$–10 Mm. These domelike separatrix surfaces enclose chromospheric small-scale loops that do not connect upward to the corona, while only a small fraction of field lines can connect upward to coronal loops along the separatrix surfaces. Essentially, the mixed-polarity magnetic field that surrounds the network field closes most of the field lines on a spatial scale comparable with the granulation (Schrijver & Title 2003; Priest et al. 2002). This scenario was simulated with detailed magnetic field extrapolations (Schrijver & Title 2002). Consequently, if we move the nanoflare scenario of Parker (1988) from the corona down to the canopy region in the transition region, we will have a much larger magnetic complexity available to drive reconnection-driven plasma heating. In coronal holes there is a larger amount of unbalanced magnetic flux, which causes more field lines to close over larger distances, which affects the channeling of upward heat flux.

3.8. *The Altitude of Nanoflares*

The most convincing argument that plasma heating resulting from nanoflares and microflare events occurs in the transition region, rather than throughout the corona in higher altitudes, are the imaging observations of EUV nanoflares and soft X-ray transient brightenings, which invariably show small-scale loops with length scales of $L \approx 2000$–5000 km that barely stick out of the transition region (e.g., Aschwanden & Parnell 2002). In the model of Parker (1988), where coronal field lines are more or less uniformly twisted along their lengths, heated plasma would spread along the coronal field lines and should manifest itself in long (probably unresolved) thin threads, for which we have no observational evidence. Furthermore, transient events such as explosive events following magnetic reconnection (e.g., Innes et al. 1997) or transition region blinkers (Harrison 1997) are seen almost exclusively in emission lines formed below some 0.3 MK.

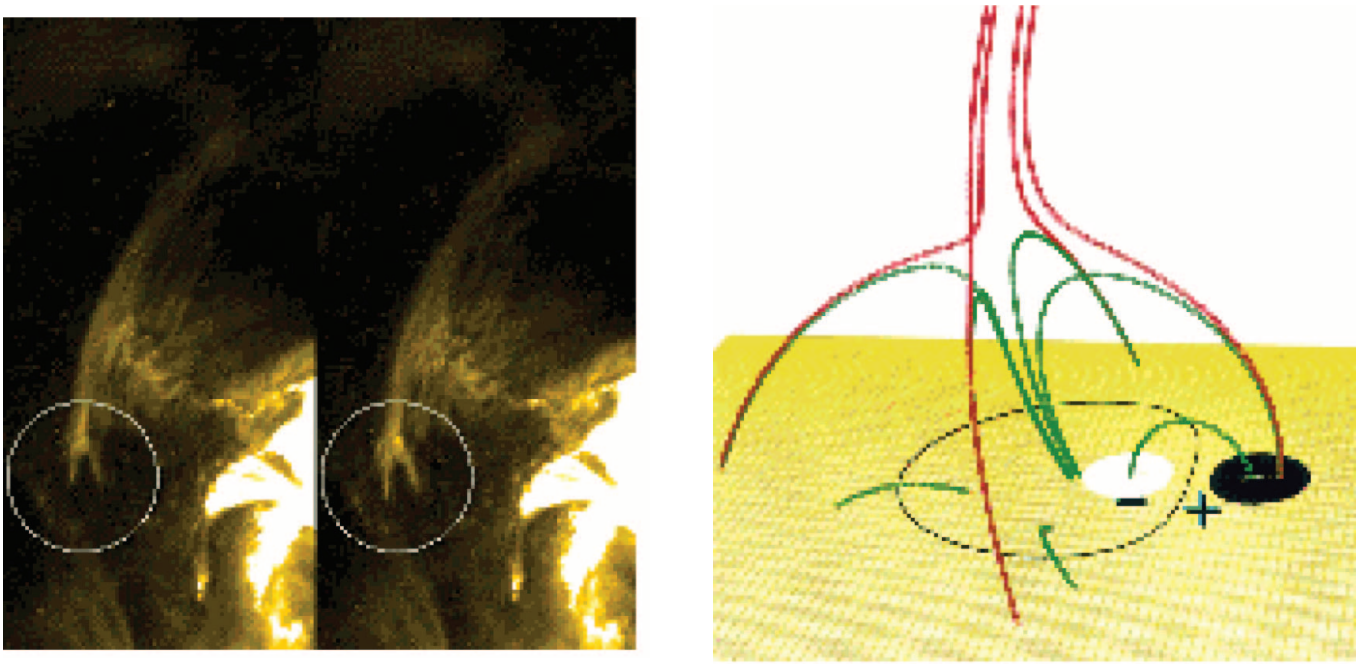

FIG. 5.—*Left:* Footpoint detail of a coronal loop recorded with *TRACE* in 195 Å ($T \approx 1.5$ MK). *Right:* Possible interpretation of the magnetic topology: An isolated magnetic polarity that is surrounded by opposite magnetic polarity forms a dome surface with a 3D magnetic null point (Brown & Priest 2001; Antiochos 1998).

This points to a high efficiency of the heating in cool rather low-lying structures.

3.9. *Elementary Loop Cross Sections*

The loop width should in principle reveal the geometric cross section of a local coronal heating event. Cross-field diffusion is strongly inhibited in the corona because of the high electric conductivity σ. Therefore, the observed loop width is not broadened by cross-field transport processes, but rather reveals the intrinsic cross section of an individual heating event. (For alternative views, e.g., see Galloway et al. [2006], where significant magnetic field fluctuations are assumed between adjacent field lines so that Rechester-Rosenbluth diffusion becomes efficient, or see also the nonresonant wave heating mechanism based on parallel electric fields generated in the MHD limit proposed by Tsiklauri [2006].) Many coronal nanoflare heating models (e.g., Parker 1988) predict unresolved small heating events. A natural consequence of this prediction is that we would observe loops with broad temperature distributions only, because every macroscopic loop structure consists then of a bundle of elementary unresolved loops. This expectation is, however, in conflict with recent results from triple-filter analysis of the finest coronal loops analyzed in *TRACE* images, where elementary loop strands with isothermal cross sections of $w \approx 1000$–2000 km were found (Aschwanden & Nightingale 2005). This observational result is hard to reconcile with any nanoflare model that predicts unresolved loop strands. However, if we move the heating process into the transition region or upper chromosphere, where the plasma β-parameter becomes comparable with unity (and exceeds unity in the chromosphere), larger heating cross sections are possible due to the lower magnetic pressure. The thermal pressure in a magnetic diffusion region in the transition region is sufficiently high to balance the magnetic pressure outside the reconnection region, and thus can lead to larger cross sections

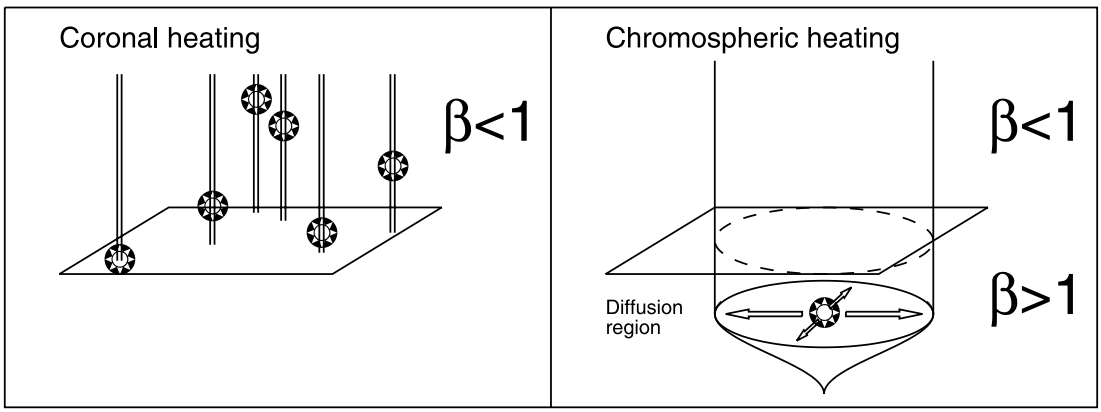

FIG. 6.—Comparison of theoretical expectations from two different heating scenarios: coronal heating (*left*) and chromospheric heating (*right*). The two scenarios predict different loop cross sections: Coronal nanoflares predict unresolved fine loop strands due to the small value of the plasma β-parameter in the corona (*left*), while chromospheric heating scenarios predict loop cross sections commensurate with the width of magnetic canopies above granulation and network cells, due to the high value of the plasma β-parameter in the chromosphere (*right*).

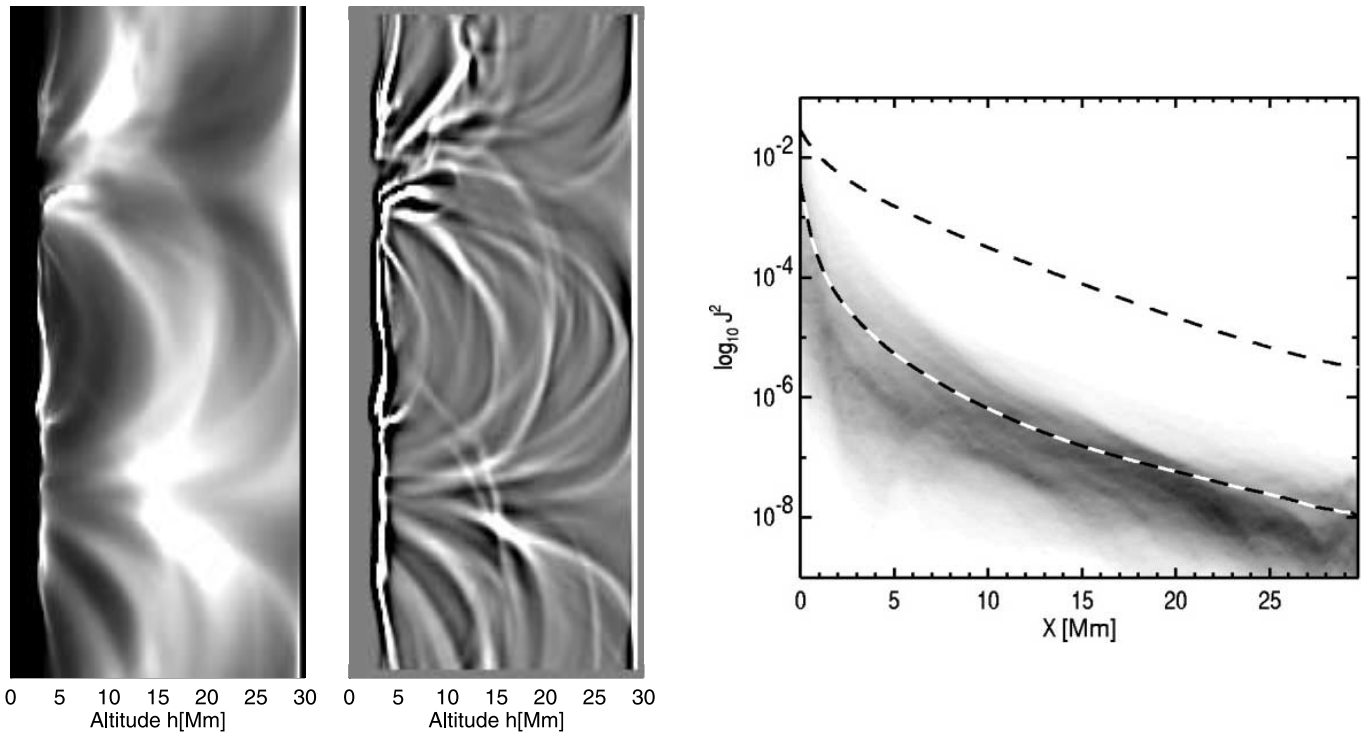

FIG. 7.—*Left:* 171 Å image of MHD simulation by Gudiksen & Nordlund (2002). The x-axis is the altitude h with the chromosphere on the left side. *Middle:* High-pass-filtered 171 Å image of simulation. *Right:* Histogram of squared current density $j^2(h)$ as a function of altitude h (marked on the x-axis with X[Mm]). Dark colors indicate a higher filling factor. The upper and lower dashed curves are the horizontal average of the squared magnetic field strength $B^2(h)$ and the squared electric current $j^2(h)$, respectively.

of loops filled with heated plasma (Fig. 6), most likely to be on scales of the granulation and network pattern, i.e., $w \gtrsim 1000$ km as observed. Thus, the observed resolved cross sections of elementary coronal loops favor chromospheric heating rather than coronal heating.

3.10. *3D MHD Simulations of Coronal Heating*

Loop cross sections of the same size as the observed elementary loops measured with *TRACE* ($w \approx 1000$–2000 km) have also been reproduced with realistic 3D MHD simulations that take the transition region structure into account (e.g., Gudiksen & Nordlund 2005a, 2005b). We show a highpass-filtered image of such a simulation in Figure 7, which displays a number of coronal loop structures with widths of $w \gtrsim 2000$ km, which are significantly wider than the horizontal pixel size of 400 km of the numerical simulations. If the nanoflare heating with unresolved strands would occur throughout the corona, isothermal loop structures should be unresolved and essentially have a width of 1 pixel.

The 3D MHD simulations reproduced also the overdensity of coronal loops: After the simulation has run for a while, the loops that are bright in *TRACE* 171 Å were on average twice as dense as the hydrostatic background loops at the same temperature ($T \approx 1$ MK).

The 3D MHD simulations yield the altitude distribution of dissipated currents, which are a measure of the coronal heating rate. The height distribution of dissipated currents falls off steeply with altitude (Fig. 7, *right panel*), about by 2 orders of magnitude in the lowest 2000 km above the photosphere, which is about the extent of the chromosphere. While the heat generated in the lower chromosphere is absorbed by strong radiative loss, some excess heating occurs in the upper chromosphere and lower transition region, which drives chromospheric evaporation and gives rise to filled coronal loops. Thus, inclusion of the chromosphere and transition region in these 3D MHD simulations of footpoint-driven braiding of coronal magnetic field lines yields the bulk of heating in the chromosphere/transition region (e.g., Gudiksen & Nordlund 2005a, 2005b). In contrast, uniform heating throughout the corona is only seen in MHD simulations that capture a coronal box (e.g., Galsgaard et al. 1999), and is therefore an artifact that results from omitting the transition region in the simulated box.

4. CONCLUSIONS

We have enumerated 10 arguments that the primary heating process leading to the appearance of hot ($T \gtrsim 1.0$ MK) coronal loops takes place in the chromosphere rather than in the corona. Most of the arguments are based on observational constraints that are often ignored in theoretical models: (1) the absence of observed temperature increases in coronal loops; (2) the overdensity of coronal loops with regard to the ambient background plasma; (3) observed upflows in coronal loops; (4) the blueshifted upflows at hot temperatures ($T \gtrsim 0.5$ MK) and redshifted downflows at cooler temperatures; (5) upward propagating acoustic waves that indicate impulsive energy releases in the chromosphere; (6) the energy balance of near-isothermal coronal loops, which can naturally be explained by hot upflows; (7) the magnetic complexity in the transition region, which is more prone to reconnection processes than the topologically simpler magnetic fields in the corona; (8) the low altitude of observed nanoflares; (9) the resolved width of elementary loop strands (based on triple-filter analysis), which require a diffusion region in a high-beta region such as the chromosphere; and (10) the reproduction of finite loop widths and preferential current dissipation in the chromosphere/transition region by 3D MHD numerical simulations of random footpoint motion-driven braiding of coronal magnetic field lines.

Given these 10 arguments, we see strong support for the working hypothesis that the major heating source responsible for a hot corona is located in the transition region rather than in the corona. We therefore consider the term "coronal heating" as paradoxical, since the essential heating process does not take place in the corona, but rather in the transition region or upper chromosphere. We suggest that the coronal heating problem should rather be addressed as "chromospheric heating problem," while the existence of a hot corona is merely a consequence of chromospheric evaporation and filling of coronal flux tubes, similar to postflare loops, but with lower densities and temperatures. This paradigm shift of the heating source from the corona to the chromosphere does not, however, exclude that the primary energy release could occur in the corona, if the coupling to the chromospheric heating is accomplished by other means than transport of heated plasma, e.g., by nonthermal particles as in flares. The new paradigm of chromospheric heating, however, does exclude models that heat active region loops directly at coronal altitudes, such as wave heating models. (Note, however, that our arguments mostly apply to the heating of active region loops, but not necessarily to open-field regions such as coronal holes; see the Appendix.)

We appreciate the very helpful and constructive comments from the referee, which helped to clarify the differences of the coronal heating problem in active regions versus coronal holes. We thank Boris Gudiksen and Aake Nordlund for providing digital data of their numerical simulations. Part of this work was supported by NASA contract NAS5-38099 (*TRACE* mission).

APPENDIX

HEATING IN CORONAL HOLES

We have to add a disclaimer here that the "coronal heating paradox" discussed in this paper applies mainly to the parts of the corona that are bright in soft X-rays and EUV, which by definition means where most of the mass is, since the flux or emission measure in (optically thin) soft X-rays and EUV scales with the squared electron density. Therefore, our arguments mostly apply to active regions and to the quiet Sun, but not necessarily to coronal holes. The physics of coronal heating can be split up into closed magnetic field regions (active regions, quiet Sun) and open-field regions (coronal holes), and since most of the 10 arguments brought forward here are loop-specific, the arguments are mostly limited to closed-field regions. Alternatively, one could split up the physics of coronal heating according to different regimes of spatial scales (and the resulting travel times and gravitational stratification), an approach that was used to discriminate between "basal coronal heating"and "extended coronal heating" (see recent reviews by Cranmer 2002, 2004; Hollweg 2006), which allows one to address different heating processes in coronal holes and in the solar wind. The physical conditions in coronal holes and in the fast solar wind are quite different and *SOHO* UVCS has revealed surprisingly large temperatures, outflow speeds, and velocity distribution anisotropies for positive ions in coronal holes. The role of Alfvén waves (their generation, propagation, and reflections) become more important on larger spatial scales (since Alfvén waves have a long dissipation length of several solar radii). Wave-particle interactions, such as ion-cyclotron resonance, are considered now as the principal mechanism for heating the coronal holes, and ultimately driving the fast solar wind (Hollweg 2006). The current understanding is that the solar wind is mainly driven by the pressure of hot protons, so the heating in coronal holes goes more into protons than electrons, because it is conveyed by the ion-cyclotron resonance rather than by currents, which is different from the DC heating models generally applied in the lower corona (basal coronal heating). UVCS observations also show us how the efficiency of heating protons and ions (e.g., O^{+5}) increases out to several solar radii, which clearly supports *extended* heating processes, which are different from the *basal* heating processes driven by photospheric random motion. In summary, there is much evidence that different heating mechanisms are dominant in the lower and extended corona (which are almost synonyms for closed-field and open-field corona). The paradigm shift emphasized in this paper can be rephrased as an argument that basal heating processes are predominant for bright parts of the corona visible in soft X-rays and EUV, which excludes the darker regions of coronal holes and the heliospheric solar wind.

REFERENCES

Antiochos, S. K. 1998, ApJ, 502, L181
Aschwanden, M. J. 1999, Sol. Phys., 190, 233
———. 2001, ApJ, 560, 1035
Aschwanden, M. J., & Acton, L. W. 2001, ApJ, 550, 475
Aschwanden, M. J., & Benz, A. O. 1997, ApJ, 480, 825
Aschwanden, M. J., & Nightingale, R. W. 2005, ApJ, 633, 499
Aschwanden, M. J., Nightingale, R. W., & Alexander, D. 2000a, ApJ, 541, 1059
Aschwanden, M. J., & Parnell, C. E. 2002a, ApJ, 572, 1048
Aschwanden, M. J., & Schrijver, C. J. 2002b, ApJS, 142, 269
Aschwanden, M. J., Tarbell, T., Nightingale, R., Schrijver, C. J., Title, A., Kankelborg, C. C., Martens, P. C. H., & Warren, H. P. 2000b, ApJ, 535, 1047
Aschwanden, M. J., Stern, R. A., & Güdel, M. 2006, ApJ, submitted
Brown, D. S., & Priest, E. R. 2001, A&A, 367, 339
Brown, J. C., Krucker, S., Guedel, M., & Benz, A. O. 2000, A&A, 359, 1185
Cranmer, S. R. 2002, Space Sci. Rev., 101, 229
———. 2004, in Proc. *SOHO-15*, Coronal Heating (Noordwijk: ESA), ed. R. W. Walsh, J. Ireland, & B. Fleck (ESA Sp-575; Paris: ESA), 154
Czaykowska, A., DePontieu, B., Alexander, D., & Rank, G. 1999, ApJ, 521, L75
DeMoortel, I., Hood, A. W., Ireland, J., & Walsh, R. W. 2002a, Sol. Phys., 209, 89
DeMoortel, I., Ireland, J., Hood, A. W., & Walsh, R. W. 2002b, A&A, 387, L13
DeMoortel, I., Ireland, J., & Walsh, R. W. 2000, A&A, 355, L23
DeMoortel, I., Ireland, J., Walsh, R. W., & Hood, A. W. 2002c, Sol. Phys., 209, 61
DePontieu, B., Tarbell, T., & Erdélyi, R. 2003, ApJ, 590, 502
Fisher, G. H., Canfield, R. C., & McClymont, A. N. 1984, ApJ, 281, L79
———. 1985a, ApJ, 289, 414
———. 1985b, ApJ, 289, 425
Galloway, R. K., Helander, P., M., & MacKinnon, A. L. 2006, ApJ, 646, 615
Galsgaard, K., Mackay, D. H., Priest, E. R., & Nordlund, A. 1999, Sol. Phys., 189, 95
Gudiksen, B. V., & Nordlund, A. A. 2002, preprint (astro-ph/0203167)
———. 2005a, ApJ, 618, 1020
———. 2005b, ApJ, 618, 1031
Harrison, R. A. 1997, Sol. Phys., 175, 467
Hollweg, J. V. 2006, Phil. Trans. R. Soc. A, 364, 505
Hori, K., Yokoyama, T., Kosugi, T., & Shibata, K. 1997, ApJ, 489, 426
———. 1998, ApJ, 500, 492
Innes, D. E. Inhester, B., Axford, W. I., & Wilhelm, K. 1997, Nature, 386, 811
Jendersie, S., & Peter, H. 2006, A&A, 460, 901
Kano, R., & Tsuneta, S. 1995, ApJ, 454, 934
MacNeice, P., McWhirter, R. W. P., Spicer, D. S., & Burgess, A. 1984, Sol. Phys., 90, 357
Mariska, J. T., & Poland, A. I. 1985, Sol. Phys., 96, 317
Müller, D. A. N., de Groof, A., Hansteen, V. H., & Peter, H. 2005, A&A, 436, 1067
Müller, D. A. N., Hansteen, V. H., & Peter, H. 2003, A&A, 411, 605
Müller, D. A. N., Peter, H., Hansteen, V. H., & Peter, H. 2004, A&A, 424, 289

Nagai, F., & Emslie, A. G. 1984, ApJ, 279, 896
Pallavicini, R., Serio, S., & Vaiana, G. S. 1977, ApJ, 216, 108
Parker, E. N. 1988, ApJ, 330, 474
Peter, H. 1999, ApJ, 516, 490
Peter, H., Gudiksen, B., & Nordlund, A. 2004, ApJ, 617, L85
———. 2006, ApJ, 638, 1086
Peter, H., & Judge, P. G. 1999, ApJ, 522, 1148
Petrie, G. J. D., Gontikakis, C., Dara, H., Tsinganos, K., & Aschwanden, M. J. 2003, A&A, 409, 1065
Priest, E. R., Heyvaerts, J. F., & Title, A. M. 2002, ApJ, 576, 533
Rosner, R., Tucker, W. H., & Vaiana, G. S. 1978, ApJ, 220, 643
Schrijver, C. J. 2001, Sol. Phys., 198, 325
Schrijver, C. J., & Title, A. M. 2002, Sol. Phys., 207, 223
———. 2003, ApJ, 597, L165
Serio, S., Peres, G., Vaiana, G. S., Golub, L., & Rosner, R. 1981, ApJ, 243, 288
Shimizu, T. 1997, Ph.D. thesis, Natl. Astron. Obs., Mitaka, Tokyo, Japan
Tsiklauri, D. 2006, A&A, 455, 1073
Tsiklauri, D., Aschwanden, M. J., Nakariakov, V. M., & Arber, T. D. 2004, A&A, 419, 1149
Tsiklauri, D., & Nakariakov, V. M. 2001, A&A, 379, 1106
Ugarte-Urra, I., Winebarger, A., & Warren, H. P. 2006, ApJ, 643, 1245
Warren, H. P., Winebarger, A. R., & Hamilton, P. S. 2002, ApJ, 579, L41
Warren, H. P., Winebarger, A. R., & Mariska, J. T. 2003, ApJ, 593, 1174
Winebarger, A. R., DeLuca, E. E., & Golub, L. 2001, ApJ, 553, L81
Winebarger, A. R., & Warren, H. P. 2005, ApJ, 626, 543
Withbroe, G. L., & Noyes, R. W. 1977, ARA&A, 15, 363
Yokoyama, T., & Shibata, K. 1998, ApJ, 494, L113
———. 2001, ApJ, 549, 1160