Solar Physics
DOI: 10.1007/•••••-•••-•••-••••-•

Resolving the 180° Ambiguity in Solar Vector Magnetic Field Data: Evaluating the Effects of Noise, Spatial Resolution, and Method Assumptions

K.D. Leka[1] · **Graham Barnes**[1] · **A.D. Crouch**[1] · **Thomas R. Metcalf**[1,2] · **G. Allen Gary**[3] · **Ju Jing**[4] · **Y. Liu**[5] ·

Received: / Accepted: / Published online:

Abstract

The objective testing of algorithms for performing ambiguity resolution in vector magnetic field data is continued, with an examination of the effects of noise in the data. Through the use of analytic magnetic field models, two types of noise are "added" prior to resolving: noise to simulate Poisson photon noise in the observed polarization spectra, and a spatial binning to simulate the effects of unresolved structure. The results are compared through the use of quantitative metrics and performance maps. We find that while no algorithm severely propagates the effects of Poisson noise beyond very local influences, some algorithms are more robust against high photon-noise levels than others. In the case of limited spatial resolution, loss of information regarding fine-scale structure can easily result in erroneous solutions. Our tests imply that photon noise and limited spatial resolution can act so as to make assumptions used in some ambiguity resolution algorithms no longer consistent with the observed magnetogram. We confirm a finding of the earlier comparison study, that results can be very sensitive to the details of the treatment of the observed boundary and the assumptions governing that treatment. We discuss the implications of these findings, given the relative sensitivities of the algorithms to the two sources

[1] NorthWest Research Associates, Colorado Research Associates Division, 3380 Mitchell Lane, Boulder, CO 80301, USA
email: leka@cora.nwra.com email: graham@cora.nwra.com
email: ash@cora.nwra.com
[2] Deceased, July 2007.
[3] Center for Space Plasma and Aeronomic Research, The University of Alabama in Huntsville, 320 Sparkman Dr., Huntsville, AL 35899
[4] New Jersey Institute of Technology, Center for Solar-Terrestrial Research 323 Martin Luther King Boulevard Newark, NJ 07102 USA
[5] Stanford University, HEPL Annex, B210 Stanford, CA 94305-4085 USA

of noise tested here. We also touch on further implications for interpreting observational vector magnetic field data for general solar physics research.

Keywords: Sun: magnetic field, Sun: observations, polarimetry, data analysis

1. Introduction

Continuing in the search for the best algorithm with which to resolve the inherent ambiguity in the transverse field component of vector magnetic field data, we present here an examination of the effects of noise in the data and imperfect instrumental spatial resolution. We build on the work of the first workshop and its resulting manuscript (Metcalf *et al.*, 2006) (hereafter "Paper I"), by presenting the results of a second workshop held at NWRA/CoRA Division in October 2006 which focused on more "real-world" challenges of observational data.

As such, we simply reiterate that without the resolution of this ambiguity which stems from a degeneracy in the linear polarization of the Zeeman effect, the physical interpretation of vector magnetic field data is of limited use. Observations at a single height do not contain sufficient information to lift this degeneracy. Thus, algorithms are based on a variety of physics-based "best guesses" or approximations implemented through a variety of optimization approaches. For a detailed description of the algorithms tested see Paper I and references therein.

In this series of comparisons a "hare and hound" approach is taken, constructing synthetic data for which an answer is known and to which the solutions arising from the application of an algorithm can be quantitatively evaluated. In Paper I, the model data tested two distinct aspects of most ambiguity resolution algorithms: first, mixed field morphologies (using a field-of-view which contained both a potential and significantly non-potential field), and second the difficulties of off-disk-center observations (using a constant-α force-free field which was viewed from an angle of just over 45°). From that first study it was realized that acute-angle azimuth resolution based on a potential-field comparison could have quite different results depending on the implementation of the potential-field calculation. It was also demonstrated that automated algorithms which globally optimized a combination of the divergence and inferred currents, performed best. Full details are given in Paper I; the most obvious unrealistic character of the model fields used therein was their smoothness and lack of noise.

A second workshop was held at NWRA/CoRA Division in October 2006 for which new model data were constructed to test the effects of Poisson noise in the observed spectra (Section 3.1), and the effects of limited instrumental spatial resolution (Section 3.2). These two aspects of observational reality were tested independently, in the hopes of separately evaluating the observational effects on this necessary part of the data reduction. Additional effort has been put into constructing appropriate evaluation metrics, as described in Section 4.1.

2. The Hounds

All participants who provided solutions in the first workshop were invited to the second, and additional interested scientists attended from the local area and from the National Solar Observatory. In the end, the provided solutions were more numerous for the first noise test (photon noise), and for the second test (spatial resolution) essentially only the best-performing automated algorithms are compared, along with different realizations of potential-field acute-angle ambiguity-resolution methods.

We refer to the methods by the names and acronyms used in Paper I, and summarize the salient differences in Table 1. In contrast to Paper I, the Non-Potential Field Calculation method ("NPFC2") (Georgoulis, 2005; Metcalf *et al.*, 2006)[1], was run by G. Barnes for this investigation. The methods tend to either locally or globally minimize a functional which may or may not depend on a model field (such as a potential or linear force-free construct). Improvements or changes in the algorithms which have been made since Paper I are described here.

2.1. The Minimum Energy Method

The Minimum Energy method (Metcalf, 1994; Metcalf *et al.*, 2006) was a work-in-progress by our late colleague Dr. Thomas R. Metcalf, and additional options and features had been recently added or were being tested at the time of his death. Included in these were approaches to "tile" the data, meaning to break a magnetogram into small sub-areas with which to compute a best-fit force-free α-parameter, effectively constructing a non-linear force-free field from which to infer the vertical gradients needed to compute the divergence. The optimization algorithm and weightings used for the functional of divergence and current were also under development. Application of the latest version of this algorithm and its complex array of options (primarily related to functional weighting, optimization method and schedule, and derivative calculations), was performed by K. D. Leka and the results are presented with the abbreviation "ME2".

We also include a version of the Minimum Energy method which is closer to the original version described in Metcalf (1994). This version, referred to as "ME0", has been recoded in Fortran by Ashley Crouch and Graham Barnes, and extensively optimized for use in the SDO/HMI data reduction pipeline[2]. It can now resolve the ambiguity in moderate-sized fields of view (of order 512^2 pixels) in roughly ten minutes. The main features of the code are as follows. The quantity to be minimized is $\lambda|J_z| + |\boldsymbol{\nabla} \cdot \mathbf{B}|$. The optional weighting factor λ is unity by default (and as implemented here). The vertical derivative needed in the approximation for the divergence of the field is obtained from a potential field calculated directly from the line of sight component of the field; thus the potential field need only be calculated once per magnetogram because it does

[1]Code is available at
`http://sd-www.jhuapl.edu/FlareGenesis/Team/Manolis/codes/ambiguity_resolution/`

[2]Code is available at `www.cora.nwra.com/AMBIG`

Table 1. Summary of tested algorithms.

Name	Quantity minimized†	Minimization scheme
Acute angle to reference field *e.g.,* Potential-Field	$\lvert\theta_{\rm o} - \theta_{\rm r}\rvert$	local
Uniform Shear Method ("USM")	$\lvert\theta_{\rm o} - \theta_{\rm r} - \Delta\theta_{mp}\rvert$	local
Non-Potential Field Calculation II ("NPFC2")	$\lvert\theta_{\rm o} - \theta_{\rm r}\rvert$, $\mathbf{B}_{\rm r} = \mathbf{B}_{\rm p} + \mathbf{B}_{\rm c}$	iterative + smoothing
Pseudo-Current ("PCM")	$\int d^2a\, J_z^2$	conjugate gradient
U. Hawai'i Iterative Method ("UHIM")	$\int d^2a\, J_z^2$, $\lvert\theta_{\rm o} - \theta_{\rm r}\rvert$	iterative
Minimum Energy v0 ("ME0")	$(\lambda\lvert J_z\rvert + \lvert\nabla\cdot\mathbf{B}\rvert)$	simulated annealing
Minimum Energy v2 ("ME2")	weighted combinations of $\lvert\mathbf{J}\rvert$, $\lvert\nabla\cdot B\rvert$	simulated annealing, genetic crossover
NCAR/HAO "`AZAM`"	angle between neighboring pixels	interactive

†: $\theta_{\rm o}$ refers to the observed azimuthal angle,
$\theta_{\rm r}$ refers to the reference-field azimuthal angle,
$\Delta\theta_{mp}$ refers to the most probable shear angle, see Paper I.

not depend on the ambiguity resolution. This removes the need for interpolation and for iteration to calculate the value of the force-free parameter to use in a linear force-free field.

To find the permutation of azimuthal angles that corresponds to the minimum of $E = \lambda|J_z| + |\boldsymbol{\nabla}\cdot\mathbf{B}|$ we use simulated annealing as described in Crouch, Barnes, and Leka (2009), except here we use a "temperature" that varies from pixel to pixel. The initial temperature is set to a value that exceeds the expected value for the change in E caused by changing the azimuth at each pixel. This modification greatly improves the efficiency over simulated annealing with a uniform temperature. In areas with large uncertainties in the transverse field, approximations to the divergence are dominated by the noise and the minimization may not return a good solution. Hence, after annealing, pixels in weak-field areas (determined with a "threshold" generally set to be below 2–3 times the noise-level of $B_\perp$) are revisited using a neighboring-pixel acute-angle algorithm similar to that used in UHIM (see Canfield *et al.* (1993)), to encourage a spatially smooth solution for noise-dominated regions. The present approach differs from UHIM in the following ways. Instead of considering pixels in order of distance from the point with the lowest continuum intensity, the pixels with the greatest number of neighbors above threshold are considered first, in order of decreasing transverse field strength. Once a pixel has been visited, it is treated as if it is above threshold (hence unlike UHIM, iteration is not needed). In determining the maximum dot product of a target pixel with its eight closest neighbors, any of the nine pixels involved which are considered to be below threshold are allowed to change their ambiguity resolution independently. Since this last step is a local approach, it is quite fast. Additional details can be found in Leka, Barnes, and Crouch (2009).

Table 2. Summary of terminology.

Term	Descriptions	Notes on azimuth
"model"/"true"	model analytic field	
	Poisson noise case:	Original analytic field
	Limited resolution case:	Binned analytic field, $\theta = \tan^{-1}(\text{binned}(B_y), \text{binned}(B_x))$
"answer"/"hare"	Field against which metrics are computed	$\lvert\Delta\theta\rvert \leq 90°$ from model
	Poisson noise:	Noise-added magnetogram's azimuth is resolved with $\lvert\Delta\theta\rvert \leq 90°$ from the model field.
	Limited nesolution:	Binned/inverted magnetogram's azimuth resolved with $\lvert\Delta\theta\rvert \leq 90°$ from the azimuth resulting from binned model field (above)
"solution"	Participants' submitted results	$\lvert\Delta\theta\rvert = 0°$or $180°$ from answer field.

3. The Hares

To perform the tests desired here required different model fields than used for the tests described in Paper I. The cases differ in how the "noise" source to be tested was applied to the model field, but the general approach to generating the test cases was the same:

1. At each pixel of an analytically-derived, spatially resolved model field, the ideal Unno-Rachkovsky Stokes profiles are generated using a Milne-Eddington atmosphere (see *e.g.,* Auer, Heasley, and House (1977), Skumanich and Lites (1987)).
2. Parameters consistent with the λ630.25nm FeI line were used: Doppler width 3.0pm (3.0×10^{-12}m, or 30mÅ), $g_{\text{landé}} = 2.5$, line-depth $\approx 0.7 I_c$.
3. To mimic observational methods, linear combinations of the pure Stokes $[I,\ Q,\ U,\ V]$ spectra were produced: $I \pm Q,\ I \pm U,\ I \pm V$.
4. These resulting "observed" profiles were convolved with an assumed 3.0pm instrumental spectral response.
5. The resulting convolved profiles were modified according to the specifics of the "noise" being tested (see below).
6. The noise-added blend spectra were demodulated to recover pure, but noise-added, Stokes $[I,\ Q,\ U,\ V]$ spectra.
7. The modified "noise-added" Stokes spectra were inverted using the same Milne-Eddington algorithm that computed the spectra from the model field originally, using standard and consistent minimization algorithms for the inversion.

In this manner, the effects of "noise" on the spectra and the resulting vector magnetic field map are isolated. The spectral inversions were initialized using the input model in order to present the inversions with the best case possible for the "initial guess" and thus avoid testing the inversion procedure itself. A

more sophisticated model for the spectral line could be used, and other common instrumental effects such as scattered light could be included, but this would add complexity to the problem of isolating the worst noise influences. Are the noise-added "observed fields" realistic? Not completely, but these cases are constructed to ensure controlled "hare and hound" tests of solely the ambiguity-resolution algorithms.

3.1. The Case of Photon Noise

Photon noise in data is an observational restriction related to the specifics of the instrumentation used; in this study we focus on this effect as it takes the form of Poisson-distributed noise in polarization spectra. The primary question we test is whether the ambiguity resolution algorithms are susceptible to propagating incorrect solutions beyond their immediate area of influence as a result of increasing noise. We specifically disregard spectral asymmetries of solar or instrumental origin, spectral distortions contributed by the blurring from atmospheric turbulence ("seeing") which are present in ground-based data, systematic polarization cross-talk, and other sources of spectral distortion.

The underlying model magnetic field is similar to that used in Paper I: a multipolar arrangement of sources was used to construct a field by means of a Green's function and a non-zero constant-α field which is force free everywhere; a simple functional, $I_c = 1 - 0.9\,|B|/\max(|B|)$ is used to calculate an effective continuum intensity. As in Paper I, one source is outside the field-of-view, hence within the field-of-view the magnetic flux is not balanced. The distribution of sources is chosen to loosely resemble an active region, includes three bald-patch areas (where the horizontal field traverses the magnetic neutral line from negative to positive polarity, indicating a concave upward magnetic field; see Titov, Priest, and Demoulin (1993)), and is computed on a regular grid in the image-plane at roughly S09 E36 ($\mu = \cos\theta \approx 0.80$, where θ is the observing angle). The general morphology of the field is shown in Figure 1.

Randomly generated noise was incorporated into the magnetograms at step #5 above, by manipulating the model Stokes spectra in the following manner. To mimic the noise in photon-counting instruments, a Poisson-distributed random number generator was used. At each spatial point, the noise-free blend spectra (from step #4) were used as the input "expectation values" to the random number generator, producing as output the "observed", noisy, blend spectra (step #5)[3]. The noise level is, according to Poisson statistics, determined by

[3]During the preparation of the manuscript, it was discovered that this protocol was not, in fact, followed for the preparation of the photon noise test cases. Co-author Metcalf had argued that polarization signals are small compared to the intensity (*i.e.* $P \ll I,\ P \in [Q,U,V]$), thus the uncertainty in $I \pm P$ is indistinguishable from that in I, and hence noisy Stokes spectra could be generated directly from the pure Stokes spectra. Using this assumption, the noisy P spectra at each point were obtained by first adding, then subtracting, two realizations of the accompanying I spectra which had been acted on by the Poisson-distributed random number generator. This differs from the correct implementation of obtaining noisy blend spectra and then demodulating (as described in the text), and resulted in the test case fields having slightly elevated noise levels. See also foootnote 4, below.

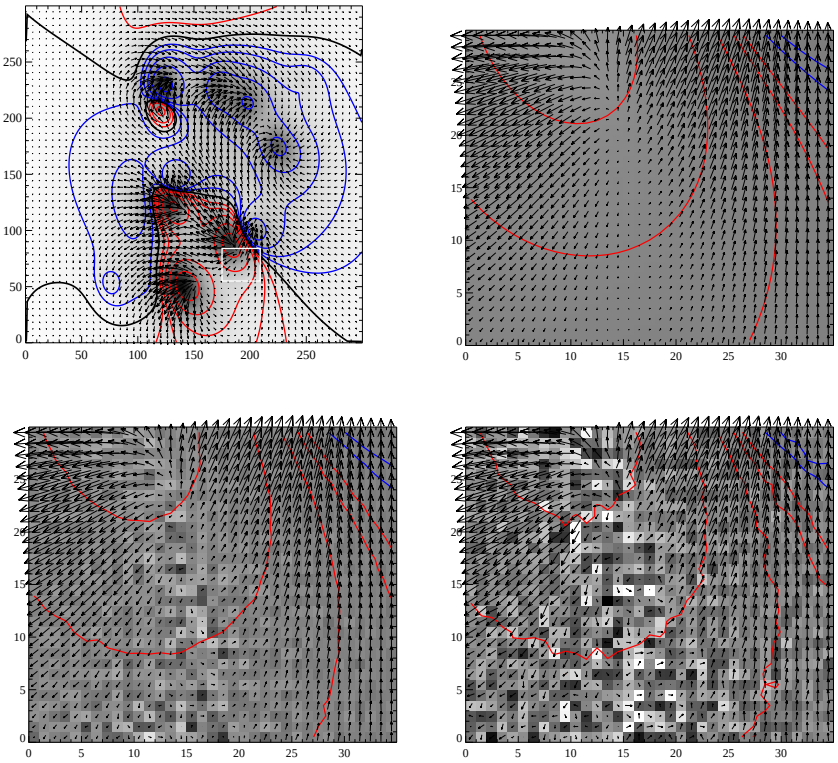

Figure 1. (Top left): the model magnetic field, for which the "observed" line-of-sight (red/blue positive/negative contours at $\pm$ 100, 200, 400, 800, 1600G) and transverse (arrows, length equivalent to one pixel $\approx$ 600G) components of the field are shown, along with the location of the *apparent* magnetic inversion line (black), all on top of a "continuum" image. (Top right): detail of the region indicated by the box on the full-area image, except the grey-scale image is the inferred vertical electric current density, scaled to $\pm$50mA m^{-2} and a 1–pixel length arrow is approximately 400G. (Bottom, left to right): detail of the low-noise and high-noise hare ("answer") fields, with the same contour levels, arrow lengths, and scaling.

the expectation value: the different degrees of noise tested were accomplished by varying the constant term of the source function (step #1).

Examples of resulting noise-added Stokes spectra are shown in Figure 2. The noise was added at two levels, roughly corresponding to $\sigma I/I_c$ levels of 10^{-2} and 10^{-3} as measured in "quiet sun" (low polarization-level regions). The range of photon noise levels considered here is consistent with data from commonly-used past, present, and future instruments such as the Imaging Vector Magnetograph ("IVM") (Mickey *et al.*, 1996; Labonte, Mickey, and Leka, 1999; Labonte, 2004), the NAOJ Solar Flare Telescope (Sakurai *et al.*, 1995), the SOLIS Vector SpectroMagnetograph (Keller and The Solis Team, 2001), the Hinode Solar Optical Telescope/SpectroPolarimeter ("SOT/SP") (Tsuneta *et al.*, 2008; Ichimoto *et al.*, 2008), and the Solar Dynamics Observatory/Helioseismic and Magnetic Imager ("HMI") (Scherrer, Hoeksema, and The HMI Team, 2006).

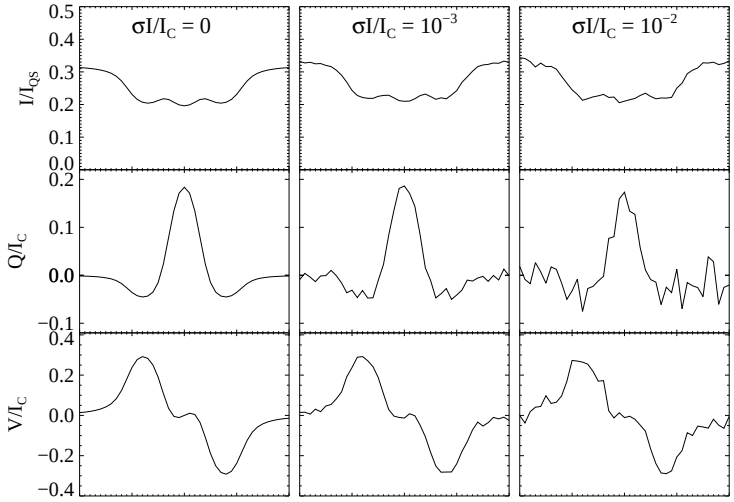

Figure 2. Examples of the addition of Poisson noise to sample spectra near one of the "umbrae". (Top to bottom): the Stokes $[I, Q, V]$ spectra normalized by the quiet-sun continuum or local continuum intensity as noted, to present the relative observed intensity level and the percent polarization, respectively. (Left to right): the three noise levels considered here, noise-free, $\sigma\ I/I_{rmc} \approx 10^{-3}$ ("low noise"), and $\sigma\ I/I_{\rm c} \approx 10^{-2}$ ("high noise").

Indeed, we build here on earlier investigations of the effects of noise in vector magnetographs, for example targeting the Marshall Space Flight Center Magnetograph (Venkatakrishnan and Gary, 1989). Initially a noise level of 10^{-4} was also considered, reminiscent of the reported sensitivity of instruments such as ZIMPOL (Stenflo and Keller, 1997), but the effect on the field was indistinguishable from the noise-free case and we do not include it here. However, to ensure consistent comparisons we also include for testing a no-noise case where the spectra have been inverted but no noise was added to the spectra. Since the continuum intensity varied as a function of field strength, there was effectively a lower signal-to-noise ratio in the "umbrae" than elsewhere, as seen in real observational data. After inverting the spectra, the effect of the noise is most obvious in the weak-field areas where the noise level and inferred field strength are comparable (Figure 1, detail panels).

The noise-added magnetograms were compared to the noise-free cases, to quantify how much "noise" was added *a priori*. The noise-added ambiguity-resolved "hare" or "answer" fields were established by imposing the acute-angle solution with reference to the original model field (see Table 2). Upon this azimuth choice, the magnitude of the angular differences between the "model" and "answer" fields were in the $-90^\circ \leq \Delta\theta < 90^\circ$ range. In other words, we allow the photon noise to impart less than 90° error in the azimuthal angle.

Two representations of the effects of this photon noise on the inferred azimuths are shown in Figure 3. The high-noise case indeed results in a greater number of pixels being more adversely affected; the areas affected are most pronounced where $\mathbf{B}_{\rm trans}$ is the smallest (whether $|\mathbf{B}|$ is intrinsically large or

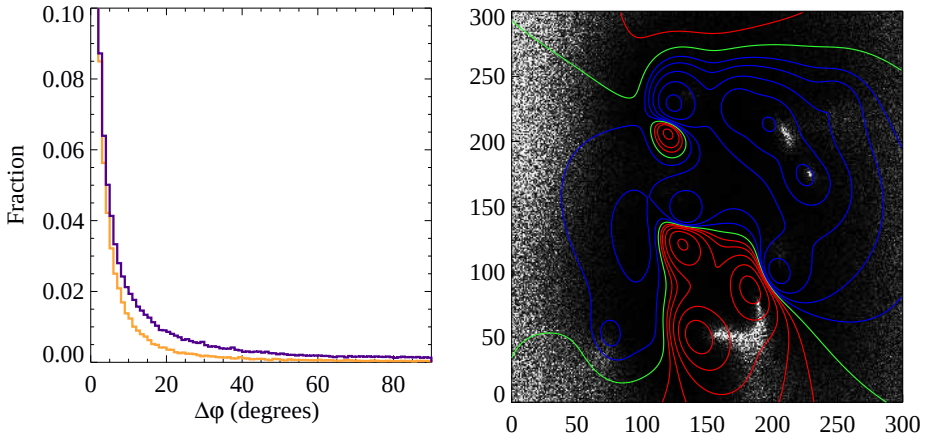

Figure 3. (Left): fractional number of points as a function of angular difference between the test cases and the original no-noise model. The former was ambiguity-resolved using an acute-angle solution with the model. Results of both added-noise levels are shown, $\sigma\ I/I_c = 10^{-3}$ "low noise" case (orange), where 68% of the points are less than 4° different, and the $\sigma\ I/I_c = 10^{-2}$ "high noise"case (purple), where 68% are less than 11° different. (Right): spatial distribution of the points with the worst angular difference, with contours of the line-of-sight field and the *apparent* magnetic inversion line following Figure 1, for the low-noise case. Grey-scale image is of the angle difference, scaled to saturate at 45°.

small). The $\mathcal{M}_{\Delta\mathbf{B}}$ metric (described in detail in Section 4.1, below) is based on a mean vector difference between a model and a test case, and can provide an estimate of the 1σ noise level imparted by the combination of added photon noise and the inversion. Here, $\mathcal{M}_{\Delta\mathbf{B}} = 0.13,\ 24,\ 71$G (gauss) for the no-noise/low-noise/high noise cases respectively when compared to the original analytic model. The largest differences introduced by the inversion for the no-noise case occur in the areas of strongest $B_{\|}$, and given the above we suggest the inversion-introduced uncertainties are not important. Similarly, the $\mathcal{M}_I$ metric quantifies the total magnitude of the vertical current of a test case, in this case vertical current introduced by the addition of the noise; for the no/low/high noise cases $\mathcal{M}_I = 1.5,\ 6.3,\ 17 \times 10^{13}$A, respectively. Inverting (the "no-noise" case) added roughly 0.1 % to $\mathcal{M}_I$ computed directly from the model field.

3.2. The Case of Limited Spatial Resolution

A second stark reality of observational data is the limit of spatial resolution. All present spectropolarimetric instrumentation and all planned for the near future is subject to the limits of spatial resolution: even at $\approx 0.1''$, there is unresolved magnetic structure on the Sun (see *e.g.,* Borrero and Solanki (2008)). The spatial mixing of spectropolarimetric signals results in an intensity-weighted average which can challenge the inversion codes, due to imposed asymmetries and combinations of $[I,\ Q,\ U,\ V]$ spectra for which a consistent atmospheric model cannot be found. The spatial mixing challenges the ambiguity resolution as what may be a smoothly-varying field on the Sun becomes discretized and

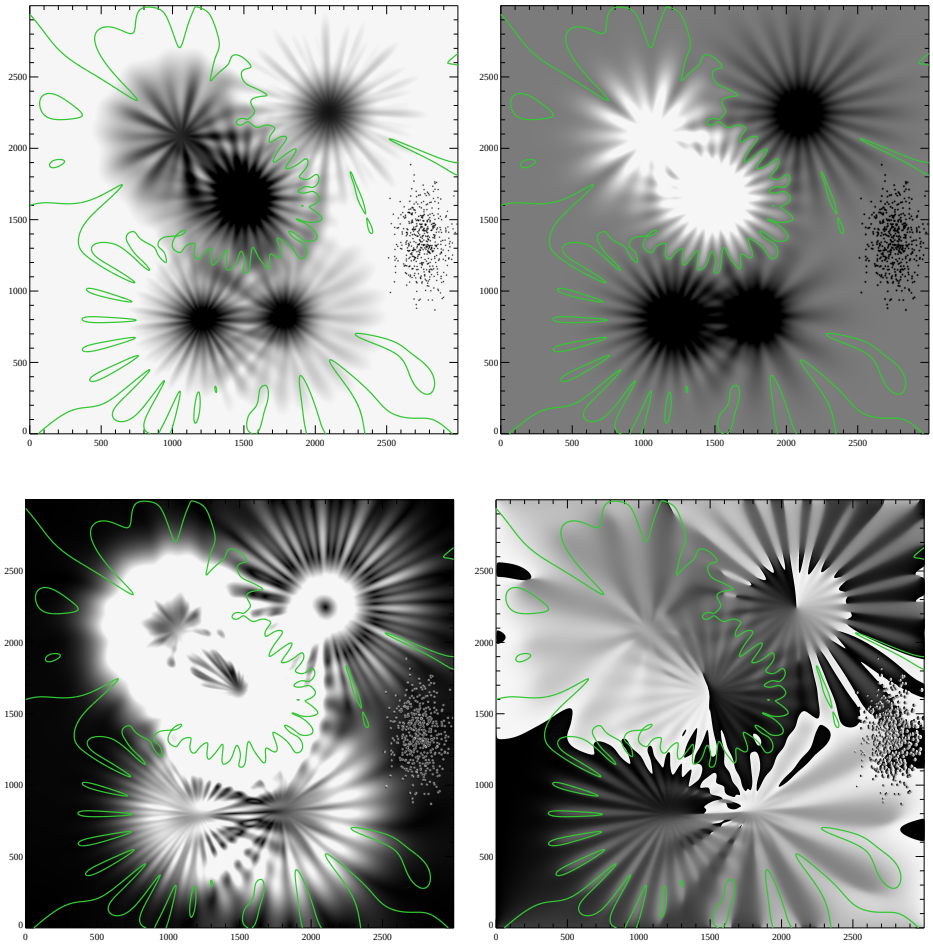

Figure 4. Full-resolution model for the "flowers" case, showing (top, left) continuum intensity (scaled from 0.2–0.9), (top, right) vertical field strength (scaled to saturate at ±1000G, white is positive), (bottom, left) horizontal field strength (scaled to saturate at 1000G), and (bottom, right) model (true) azimuth, scaled black/white to 0°, 360°. For all, the magnetic neutral line is also indicated for reference (green).

"jumpy" on the observed grid, with large pixel-to-pixel variations in strength and azimuth of the transverse field.

To test this aspect of robustness in ambiguity resolution algorithms, a model was devised so as to be current-free (potential) but possess fine-scale structure reminiscent of both sunspot penumbrae and solar plage areas.

The "flowers" test case is constructed by specifying the normal magnetic field component on two horizontal surfaces, and calculating the horizontal components of the corresponding potential field. By placing the two surfaces close together (in this case, the equivalent of six horizontal pixels on the original 0.03″ grid apart in height), it is possible to control the direction of the horizontal field

to some degree. For example, by using the same normal component on both surfaces, the field can be made more vertical than if the second surface were absent.

Each individual "sunspot" is constructed from the following normal magnetic field component on each surface

$$\begin{aligned} B_{\rm s} &= B_0 \exp(-r^2/a_0^2) + B_1(r^2/a_1^2)\exp(-r^2/a_1^2)[1+\cos(n_1\theta+\phi_1)] \\ &\quad + B_2(r^2/a_2^2)\exp(-r^2/a_2^2)[1+\cos(n_2\theta+\phi_2)] \end{aligned} \tag{1}$$

where the first term forms an axisymmetric "umbra", and the other two terms produce the pattern of "petals" similar to penumbral fibrils. The value of each of the parameters, B_j, a_j, n_j, ϕ_j is typically different on the two surfaces. By introducing a phase shift, $\Delta\phi_j$, between the petals on the upper and lower surfaces, it is possible to produce a horizontal field with a significant azimuthal component, while by changing the radii, r_j, on the two surfaces, the field can be inclined to a greater or lesser degree, in both the umbra and penumbra. However, the combination of B_j and a_j is chosen such that the net flux through the lower boundary is equal to the net flux through the upper boundary.

One can see from a detailed examination of Figure 4, that in the top/right "sunspot" at the "radius" where the total field strength falls dramatically and the fine-scale structure becomes increasingly variable azimuthally, the general orientation of the horizontal field switches direction, from being essentially radially inward (as expected from a negative-polarity sunspot) to outward. There is a ring of very small, but potential, "azimuth centers" in the underlying model. This may be dismissed as being "non-solar", but we argue that it may be appropriately descriptive of the many small structures near penumbral/moat boundaries.

The plage region (in the center/right portion of the model) is created by randomly distributing individual "flux tubes" of the same sign across a specified area. The probability distribution for the locations of the flux tubes is initially Gaussian. However, to prevent the superposition of two flux tubes, a minimum separation between the centers of the flux tubes is imposed. The normal component of each flux tube is given by a simple Gaussian

$$B_p = B_0 \exp(-r^2/a_0^2)\,, \tag{2}$$

but the values of B_0 and a_0 are allowed to vary slightly from one tube to the next. The value of a_0 is chosen such that the flux tubes are a few pixels across in the unbinned case, and thus spatially resolved. For the plage field presented here, $a_0 = 5.2$ pixels, $B_0 = -1.8$ kG, and the minimum separation was ≈ 6.5 pixels.

The continuum intensity for this model field was computed using the field strength weighted for the vertical component:

$$I_{\rm c} = \begin{cases} 1. - 0.9\,b/(2\times10^3), & \text{for } b \le 2\times10^3, \\ 0.1, & \text{otherwise,} \end{cases} \tag{3}$$

where

$$b = \sqrt{0.2\,(B_x^2 + B_y^2) + B_z^2}, \tag{4}$$

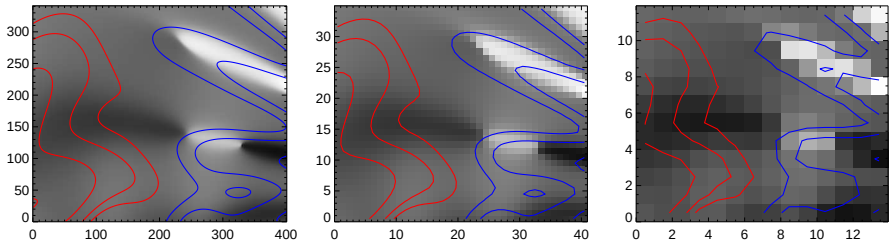

Figure 5. Images of the model-field azimuth for a small area between the top two opposite-polarity "sunspots", for the different spatial binnings tested here. Contours indicate the vertical mean magnetic field (red is positive) at 100,200,400,800 G. (Left to right): original (0.03″), bin-factor 10 (0.3″), and bin-factor 30 (0.9″). Note the lack of smoothness, implying a deviation from potential in the latter two, simply due to the spatial binning and limited sampling.

that is, a function based on the field strengths, scaled to a maximum of 2kG (the model field maximum is $\approx$ 3kG) and normalized by the same; this reproduced a fairly "solar-like" pseudo-continuum looking image, as shown in Figure 4.

This model field was computed on a fully resolved grid assumed to represent 0.03″. The emergent polarization spectra were computed as per the steps outlined above, combined to reflect observed $I \pm Q$ *etc*, spatially averaged by factors of 5, 10 and 30, demodulated to pure polarization states again, and re-inverted. Noting that there is no solar atmosphere and there are no plasma velocities to introduce asymmetric polarimetric spectra as often observed in sunspots, the present case is greatly simplified over real observational data. The resulting magnetograms have model spatial resolutions of roughly 0.15″ and 0.3″ (consistent with the resolution provided by Hinode SOT/SP magnetogram data (Tsuneta *et al.*, 2008) and high-resolution ground-based instrumentation such as the Diffraction Limited SpectroPolarimeter (Lites *et al.*, 2003; Sankarasubramanian *et al.*, 2006)), and 0.9″ (which is consistent to that achieved by instruments such as the IVM, SOLIS, and HMI on SDO).

As shown in Figure 5, there is small-scale structure which is smoothly varying on the fully resolved model boundary, but becomes significantly less so with spatial binning. This results directly from the averaging of the intermediate $I \pm P$ ($P \in [Q,\ U,\ V]$) spectra over a variety of small-scale structures, each of which (after demodulating) produces its specific set of pure Stokes spectra but with an intensity-weighted averaging; the resulting Stokes spectra for any particular binned pixel may, or may not, still be consistent with the Milne-Eddington Unno-Rachkovsky approximations. This is demonstrated in Figure 6 for a selection of Stokes spectra from points in the "plage" area of the model field and a binning factor of 10. In this case, the 100-pixel area of the underlying model field contains points with field parameters that cover a broad range:

$$\begin{aligned} 76\ \mathrm{Mx\,cm^{-2}} &\leq |\mathbf{B}| \leq 2017\ \mathrm{Mx\,cm^{-2}} \\ -81^\circ &\leq \gamma \leq -11^\circ \\ -178^\circ &\leq \theta \leq +179^\circ \end{aligned}$$

whereas the resulting inverted bin-10 field at the relevant point has:

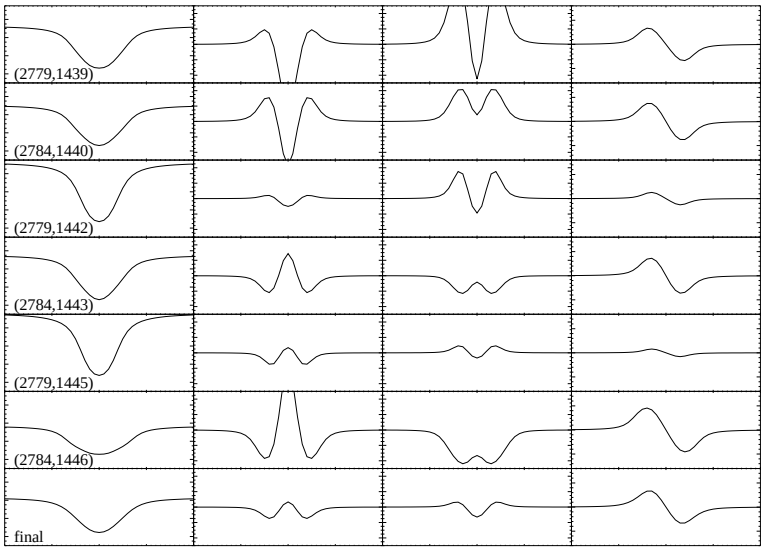

Figure 6. (Top to bottom): selection of six [I, Q, U, V] spectra from the original model field in the plage area. When binned by a factor of 10 (100 points averaged) for the 0.3″ case, the bottom row of [I, Q, U, V] spectra result. All are plotted on the same scale for their polarization type (that is, all [I] spectra are on the same scale, all [Q, U] spectra are on the same scale to each other but different from [I], and same for [V] spectra).

$$\begin{aligned} |\mathbf{B}| &= 633\ \mathrm{Mx\,cm^{-2}} \\ \gamma &= -67^\circ \\ \theta &= 69^\circ \text{ or } -111^\circ \end{aligned}$$

(with θ subject to the 180° ambiguity). This raises at least the question of whether the resulting point is an acceptable average representation of the underlying field. There is simply no way that this single binned point can reflect all of the physics of the underlying field (recall Parker (1996)), yet that is how observational data are routinely interpreted.

Again, metrics were computed for the "hare", or answer, fields (see Table 2) of different spatial resolutions, inverted and ambiguity-resolved using an acute-angle solution with reference to the directly-binned model field. A bin-1 (no binning, but inverted and acute-angle ambiguity resolved) case is included here for completeness. We find that the binning and inversion result in $\mathcal{M}_{\Delta\mathbf{B}} = 0.25,\ 1.8,\ 4.7,\ 20\mathrm{G}$ and $\mathcal{M}_I = 0.1,\ 3.4,\ 4.6,\ 7.8 \times 10^{12}\mathrm{A}$ for the $0.03''$, $0.15''$, $0.3''$, $0.9''$ cases respectively. In a small fraction of the pixels (less than 1%), the angular differences were significant between the hare and the model, indicating a strong influence of the spatial sampling (see Figure 7). Fewer than 0.01% of the points in the unbinned but inverted case resulted in a change in azimuth due solely to the inversion. Certain areas proved to be most affected by the limited resolution in this manner (Figure 7), and below we show that these areas were indeed the most troublesome for the ambiguity resolution algorithms.

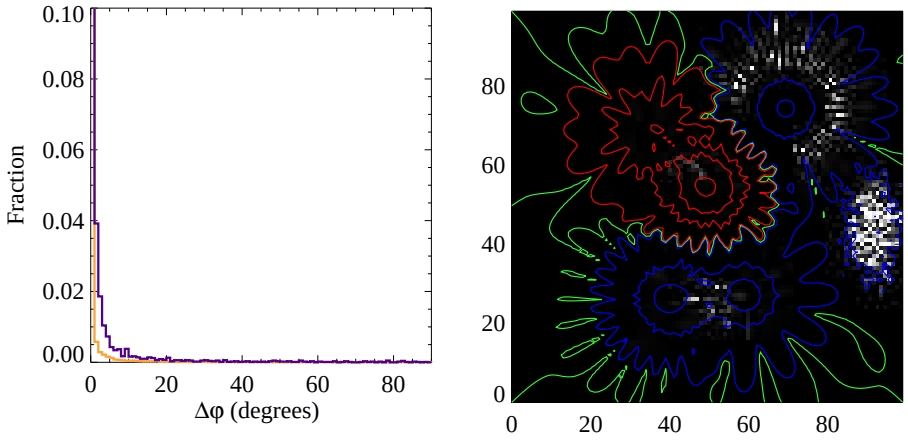

Figure 7. (Left): fractional number of points of angular difference between the azimuth resulting from a direct binning of the original model field and that from the inversion of the binned spectra. The latter was ambiguity-resolved using the former such that no difference is greater than 90°. Results of both binnings are shown, 0.3″ (orange), where 97.7% of the points are less than 1° different, and the 0.9″ (purple), where 88.0% are less than 1° different. In both cases, the 1σ level (68% of the points) falls below 0.5° difference. (Right): spatial distribution of the points with the worst angular difference for the 0.9″ case (see text for details), with contours of vertical field and magnetic neutral line shown following Figure 4. Grey-scale image is of the angle difference, scaled to saturate at 45°. The 0.3″ case shows similar degree and locales of disagreement.

With limited spatial resolution, no finite-difference calculation is guaranteed to accurately represent the spatial derivatives of the original field. The ability to infer the potential-field nature of the underlying field degrades, and the presence of vertical currents is inferred. The discretized nature of the binned boundary field impacts the inferred horizontal derivatives, and any resulting calculation which depends on them, such as $\nabla_h \cdot \mathbf{B}$ and $\nabla_h \times \mathbf{B}$.

To quantify the degree to which the lack of spatial resolution affects the field, we construct the quantity

$$\mathcal{D} \;=\; \frac{|\partial B_x/\partial x + \partial B_y/\partial y + \partial B_z/\partial z|}{3^{1/2}[(\partial B_x/\partial x)^2 + (\partial B_y/\partial y)^2 + (\partial B_z/\partial z)^2]^{1/2}}, \qquad (5)$$

in which all of the derivatives are approximated by centered finite differences. The first height is the "binned model" used as reference for the metrics below; the second model "binned model" height is treated consistently. The two heights are used to approximate the vertical derivatives (Figure 8). The normalization is such that $\mathcal{D} \leq 1.0$. For a fully resolved field, in which the derivatives are accurately represented by finite differences, $\mathcal{D} \ll 1$, since the numerator is simply $\nabla \cdot \mathbf{B}$. For ambiguity resolution methods which rely on horizontal derivatives to compute either the vertical current density or the divergence of the field, large values of $\mathcal{D}$ indicate that the value of $\mathbf{J}_z$ or $\nabla \cdot \mathbf{B}$ arrived at by finite differences will not accurately reflect the underlying field. Thus one expects that areas of large $\mathcal{D}$ may predict areas where the ambiguity resolution algorithms fail.

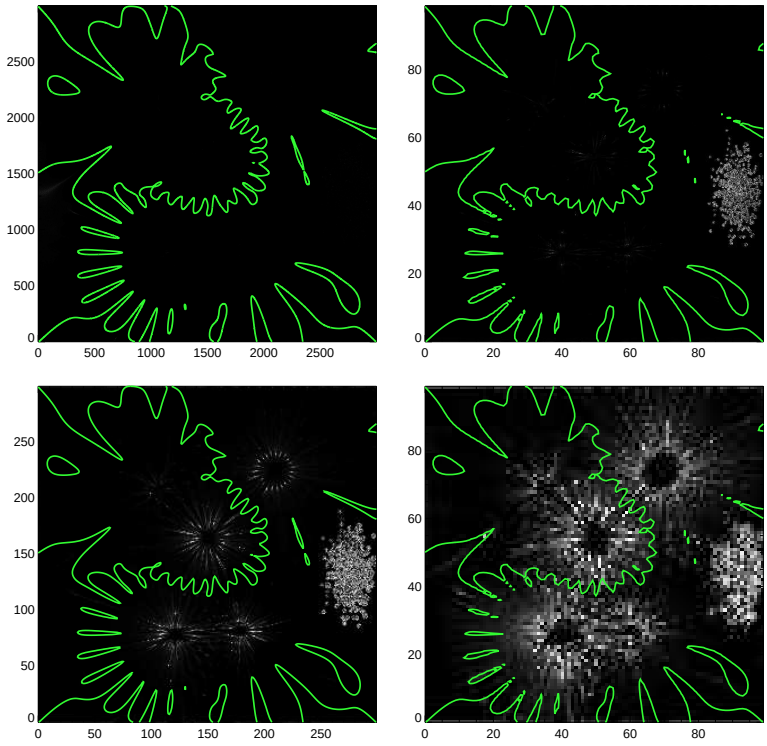

Figure 8. Images of the parameter $\mathcal{D}$, indicating (by grey-scale) where finite-differences succeed (black, $\mathcal{D} = 0$) and fail (white, $\mathcal{D} = 1$) to reproduce the underlying model field's derivatives. The contour indicates the magnetic neutral line, for reference. (Top, left to right): original ($0.03''$) fully resolved model, bin-5 ($0.15''$), (Bottom, left to right): bin-10 ($0.3''$), and bin-30 ($0.9''$).

Are the models and binning levels appropriate? To demonstrate that they are, we show in Figure 9 a comparison of magnetic flux densities in the line-of-sight and transverse components observed in plage and penumbra by Hinode SOT/SP in "fast-map" mode with $0.32''$ sampling, compared to the analogous structures resulting from the bin-10 (simulating $0.3''$ sampling). While the observational data have more scatter than the model-derived data, the general characteristics in terms of distribution functions and magnitude limits match well.

4. The Metrics and Results

In the first paper, quantitative metrics were devised to evaluate the algorithms' performances. We also presented spatial maps of success and failure, for insight into which particular structures presented the most challenge.

In the present cases, the addition of noise precludes exactly this approach. The underlying model is known (referred to as the *"true" case*), but the addition of noise and subsequent inversion of the affected Stokes spectra (resulting in

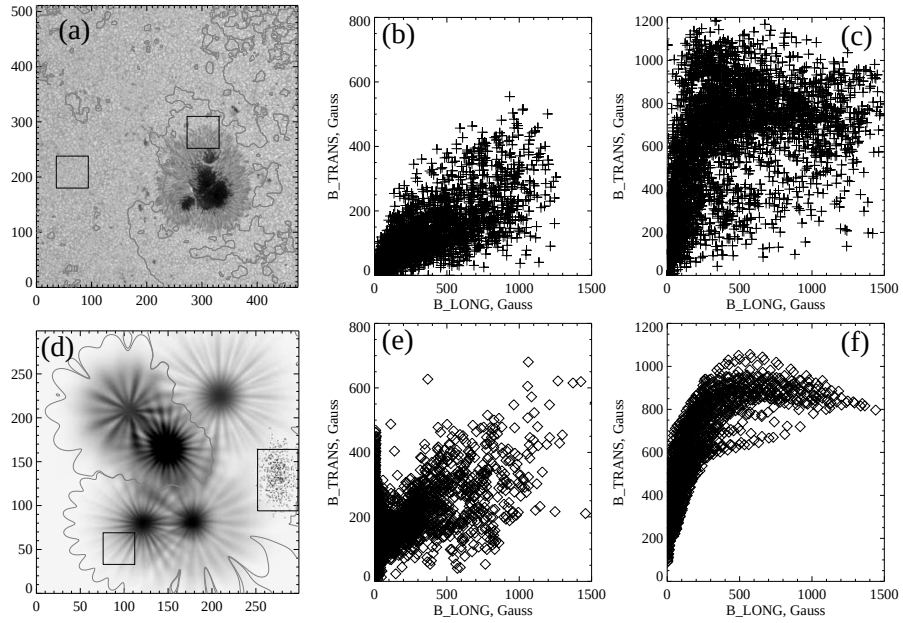

Figure 9. Comparison of the bin-10 data which approximates 0.3″ spatial resolution with observational vector-field data from the Solar Optical Telescope/Spectropolarimeter aboard the Hinode mission, from 30 April 2007. (Top row, left to right): continuum images from the SP data, with boxes indicating selected areas of "plage" and "penumbra", and the distribution of the observed line-of-sight and transverse magnetic flux densities for the "plage" area (middle), and the "penumbral" area (right). (Bottom row, left to right): same as top row, but for the model data.

ambiguous azimuthal angles), precludes direct comparison without additional levels of interpretation. We address this issue in two ways.

First, as alluded to above (Table 2), the hare, or *"answer" map* is produced using an acute-angle calculation minimizing the difference between the azimuth derived from the noise-added inversion and the original noise-free model or its proxy. In the case of the added Poisson noise, the "answer" was the direction within 90° of the noise-free model; in the limited-resolution cases, the "answer" was the direction within 90° of a proxy model field derived by directly binning the original model. By "directly binning", it is meant that (1) the model $B_x,\ B_y,\ B_z$ components are spatially averaged by the specified bin factor, (2) then $\theta = \tan^{-1}(\mathrm{binned}(B_y), \mathrm{binned}(B_x))$, resulting in ambiguity-resolved azimuths. Second, a few additional metrics were devised that were less dependent on these "answer" maps, relying instead on global quantities of the submitted solutions.

Using these hares and the participants' submitted solutions, we compute metrics and the spatial maps of success/failure in the same manner as in Paper I. Still, it must be emphasized that interpreting the performances is hampered by the fact that the presence of noise precludes knowing "the right answer", even when using synthetic data.

4.1. The Metrics

Finding from Paper I that various metrics added only limited insight, we present select metrics from the full list in Paper I, and then develop additional metrics described below which add insight for these cases. The metrics from Paper I which are used here are:

- **Area**: the simple fraction of pixels which were resolved correctly: $\mathcal{M}(\mathrm{a,s})_{\mathrm{area}} = \#\mathrm{pixels}(\Delta\theta = 0)/\#\mathrm{pixels}$. "Best" is 1.00; "random" gives 0.50.
- **Normalized Vertical Current Density**: Following paper I, the vertical currents were computed for both the submitted solutions and the "answer" map, and compared. In this metric, the solution is rewarded where it is correct *and* penalized where it is incorrect:

$$\mathcal{M}(\mathrm{a,s})_{J_z} = 1 - \frac{\sum(|J_z(\mathrm{answer}) - J_z(\mathrm{solution})|)}{2\sum(|J_z(\mathrm{answer})|)}. \tag{6}$$

  $\mathcal{M}(\mathrm{a,s})_{J_z}$ is normalized such that "good" is closer to 1.00, and 0.00 occurs if the current is exactly reversed at each pixel; scores less than 0.00 indicate the presence of strong line currents. A score of about 0.00 can also be attained with a combination of exactly reversed currents and moderate line currents.

The additional metrics described below attempt to measure sensitivity to the specific type of noise being tested.

- **Total Vertical Current**: The total unsigned vertical current, $\mathcal{M}_I = \sum \mid J_z(\mathrm{s}) \mid$, computed for the submitted solutions and also for the "answer" maps for comparison (*c.f.* Section 3.1).
- **Transverse Field**: the fraction of transverse field above a specified threshold $\mathcal{T}$ which was resolved correctly: $\mathcal{M}(\mathrm{a,s})_{B_\perp > \mathcal{T}} = \sum(B_\perp(\mathrm{s})_{\Delta\theta=0, B_\perp>\mathcal{T}})/\sum(B_\perp(\mathrm{s})_{B_\perp>\mathcal{T}})$.
- **Mean Vector Difference Magnitude:** The average magnitude of the vector difference between the solution and the model field: $\mathcal{M}_{\Delta\mathbf{B}} = \sum|\mathbf{B}(\mathrm{s}) - \mathbf{B}(\mathrm{m})|/\#\mathrm{pixels}$. In units of Gauss, this metric quantifies how much additional noise is generated by the ambiguity resolution, and is strongly weighted such that the contribution of incorrect answers in weak-field areas is less significant. Computed for the submitted solution and also for the "answer" maps as a 1σ noise level (*c.f.* Section 3.2).

For each of the metrics, a submitted image-plane azimuth solution was used to resolve the ambiguity in the test magnetogram, and the vector components of the image-plane and helioplanar field were computed from this solution. The vertical current densities were computed using finite-differences as described in Canfield *et al.* (1993). Most importantly, all quantities required for the metrics were calculated consistently. The $\mathcal{M}_I$ metric should be very discriminating for the tests of spatial resolution: the field inferred after spatially binning and then

re-inverting the spectra from the original current-free model may no longer have zero vertical current everywhere. The results of $\mathcal{M}_I$ are also presented for the "answer", as it too is affected by the addition of the spatial-binning noise and is no longer current-free as determined using a standard finite-difference approach.

The $\mathcal{M}(\mathrm{a,s})_{B_\perp > \mathcal{T}}$ metric is applied for both noise-tests, for in truth it quantifies how well an algorithm performs in the stronger-transverse-field areas where it is arguably most important (or at least where the polarization signals and hence the field is best determined). We present results for two thresholds, 100G and 500G, to highlight the performances in weak and strong field areas.

The $\mathcal{M}_{\mathbf{\Delta B}}$ metric is useful for both tests, and could be considered a "noise level" introduced by an incorrect ambiguity resolution. The results for this metric are also included for the "answer" or noise-added hares as compared to the "true" model or its proxy, for reference.

4.2. Results for Photon-Noise Cases:

The results are presented below in both quantitative metric and graphical form. We present the metrics for all submitted solutions, but only graphical results for a select few in order to highlight specific trends.

In general, the algorithms performed worse with increasing photon noise, as expected (see Table 3)[4]. The notable exception was the "Pseudo-Current" method, which performed better as noise levels increased (see Figure 10). This unexpected (but not yet fully understood) effect is possibly achieved due to the method's focus on the major current systems and minimal influence from weak magnetic field contributions (those dominated by noise), resulting in a minimal sensitivity to the presence of noise. The effects of noise are indeed the worst where $B_\perp$ signal is weak, as demonstrated by the $\mathcal{M}(\mathrm{a,s})_{B_\perp > 100\mathrm{G},500\mathrm{G}}$ metrics: seven of the methods scored perfectly for all noise levels for $\mathcal{M}(\mathrm{a,s})_{B_\perp > 500\mathrm{G}}$ whereas none achieved the same feat for $\mathcal{M}(\mathrm{a,s})_{B_\perp > 100\mathrm{G}}$. This is confirmed by comparing Figure 10 and Figure 3, the areas where the best algorithms had problems were generally areas of lowest transverse field strength, where essentially the noise dominates the signal and a good performance simply cannot be expected. This may occur in areas of weak field, but may also be the case where field strengths are large but the observing angle is such that there is minimal transverse field signal. The value of relying on multiple metrics is also clear, since even with a perfect $\mathcal{M}(\mathrm{a,s})_{B_\perp > 500\mathrm{G}} = 1.00$ score, different algorithms will score differently regarding the average field strength difference metric, $\mathcal{M}_{\mathbf{\Delta B}}$. This differentiates the UHIM results as compared to the NPFC2 results, for example. The effects of increasing noise are more or less impacting for different algorithms; some show steep changes for increasing noise levels (ME0, ME2), whereas others are relatively insensitive (NPFC2, potential). Note that for the potential field methods,

[4] A full metrics analysis was performed on maps generated by the protocol exactly described in Section 3 (see footnote #3) for the potential-field acute-angle and the UHIM algorithms. The resulting metrics differed from those shown here by at most a few percent, and there were no unexpected systematic trends. We conclude that there is no significant impact to the results of this study.

the way in which the potential field is calculated typically has a much larger impact than the changes in the noise level. These differences in implementation are similar to what was found in Paper I.

Fundamentally, the addition of photon noise and its effects on the inferred "observed" field is to make the latter inconsistent with most of the assumptions being made by most algorithms. As discussed in Paper I, most algorithms assume that observed magnetograms reflect a Sun which is in a state of minimum current and/or divergence; here, we present evidence that while the original field satisfied these criteria, the observed field no longer does. Indeed, the final global minimum of vertical current and divergence for ME0 for these cases fell *below* that for the "answer" field, yet ME0 alone (and ME2) returned decent but not great results. Similarly, when the NPFC2 algorithm was initiated with the vertical current density of the answer field, it did not correctly resolve all pixels. Instead, it resulted in a solution which scored similarly to that presented in the tables, achieved with no additional information. Thus we conclude that its assumptions are also violated by the addition of photon noise.

Algorithms which incorporate nearest-neighbor smoothing (NPFC2, AZAM) performed most robustly for this test. The recognition of the value-added for a nearest-neighbor smoothing component initiated the addition of this component to ME0, substantially improving its performance in the case of photon noise dominating the linear polarization signal.

4.3. Results for Limited-Resolution Cases:

The second test case proved challenging to most algorithms. We again present the results for this challenge in both quantitative metric (Table 4) and graphical (Figure 11) form, limiting the presentation of the graphical results to highlight a few trends. Note that not all algorithms submitted solutions to test against this source of noise.

The areas which are most affected by the spectra's spatial averaging and subsequent inversion are the plage and regions between like-polarity sunspots where the underlying model field (for example) smoothly passes through zero by transitioning from being directed $+\hat{x}$ to $-\hat{x}$, as between the two same-polarity sunspot centers. We focus on the areas where finite differences may not accurately approximate derivatives, to examine whether they especially become problematic for ambiguity resolution algorithms (Figures 8), keeping in mind the noise arising from the "mixing" of polarization signals (Figures 7).

In addition, the ring of azimuth centers that is part of the "sunspot" in the north/west corner of the field of view is constructed such that outside of this ring, the direction of the field is opposite that which one would generally infer for a "negative-polarity sunspot", with the field being directed radially outward from the center of the spot in this case. This aspect is analogous to the case in Paper I of a potential-field arcade overlying a twisted flux-tube, where the horizontal field direction at the edges was opposite what would normally be expected. When supplied with only the lower boundary "observation", all methods need additional information regarding the nature of the field, hence all methods make assumptions to furnish that missing information. As in Paper I we find many methods perform poorly when the underlying field violates the assumptions being made.

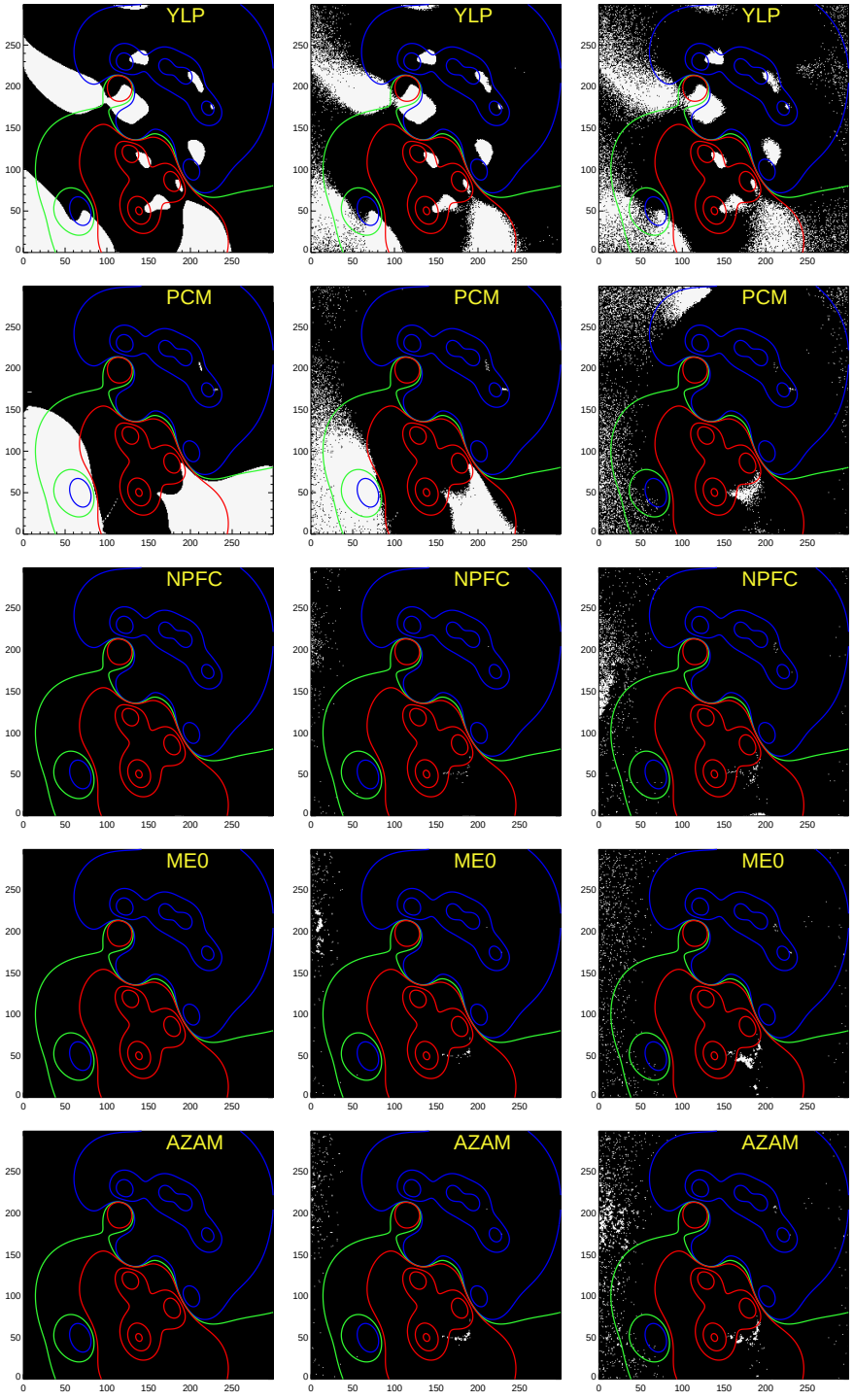

Figure 10. Examples of the effect of Poisson noise on the ambiguity resolution algorithms. Throughout, black/white areas in the underlying image represent areas which were/were not solved correctly, and red/blue contours indicate positive/negative $\mathbf{B}_z$ at $\pm$ 50, 700, 1500, 2800 G, with the true magnetic neutral line shown by a green line. (Left to right): no-noise, low-noise and high-noise cases, (from top): results for a Potential-Field Acute Angle method, for the Pseudo-Current method, for NPFC2, for ME0, and `AZAM`. Results for other potential-field and linear force-free field based solutions were similar in nature to that shown at the top: large swaths with some smaller patches which were incorrect. The general trend shown here was that for most methods: the presence of noise did not generally lead to dramatic changes in the results or new propagation of erroneous solutions, but additional "salt and pepper" of correct/incorrect solutions in weak-$B_\perp$ areas. In one case, the Pseudo-Current method (second from top), the results improved with noise.

Table 3. Results for noise-added cases.

Noise level:	$\mathcal{M}(\mathrm{a,s})_{\mathrm{area}}$			$\mathcal{M}(\mathrm{a,s})_{J_z}$			$\mathcal{M}(\mathrm{a,s})_{B_\perp>100\mathrm{G}}$			$\mathcal{M}(\mathrm{a,s})_{B_\perp>500\mathrm{G}}$			$\mathcal{M}_I$ (10^{13}A)			$\mathcal{M}_{\Delta\mathbf{B}}$		
	none	low	high	none	low	high	none	low	high	none	low	high	none	low	high	none	low	high
"Answer" field													1.53	6.27	17	0.13	24.0	71.3
Potential, FFT																		
YLP (Y. Liu)	0.81	0.81	0.81	0.30	0.52	0.57	0.91	0.91	0.89	0.99	0.98	0.98	3.4	9.0	22	79	97	139
KLP (K.D Leka)	0.64	0.65	0.65	0.53	0.41	0.44	0.85	0.84	0.82	0.98	0.99	0.98	2.8	8.1	20	139	152	187
JJP (J. Jing)	0.66	0.66	0.66	0.31	0.36	0.40	0.84	0.84	0.82	0.97	0.97	0.97	3.4	8.9	22	140	153	189
GBP (G. Barnes)	0.63	0.62	0.67	0.53	0.33	0.37	0.82	0.80	0.79	0.96	0.96	0.96	2.9	8.7	23	146	167	180
Acute Angle (LFFF)																		
HSO (H.N. Wang)	0.85	0.85	0.83	0.67	0.51	0.53	0.96	0.95	0.92	1.00	1.00	0.99	2.3	9.5	23	46	64	115
Uniform Shear Method																		
USM (Y.-J. Moon)	0.94	0.93	0.91	0.92	0.84	0.82	1.00	0.99	0.97	1.00	1.00	1.00	1.7	6.6	17	10	33	82
Non-potential Field Field Calculation II																		
NPFC2 (G. Barnes)	1.00	0.99	0.98	0.99	0.98	0.95	1.00	1.00	0.99	1.00	1.00	1.00	1.5	6.3	17	0.16	24	72
Pseudo-Current																		
PCM (A. Gary)	0.77	0.84	0.94	0.59	0.64	0.81	0.90	0.94	0.97	1.00	1.00	1.00	2.5	7.8	18	86	69	85
UH Iterative																		
UHIM (K. Leka)	0.98	0.91	0.91	0.94	0.84	0.84	0.99	0.98	0.97	1.00	1.00	1.00	1.6	6.4	17	8.2	41	86
Minimum Energy																		
ME0 (Crouch/Barnes/Leka)	1.00	1.00	0.99	1.00	0.98	0.96	1.00	1.00	0.99	1.00	1.00	1.00	1.5	6.3	17	0.13	24	72
ME2 (K.D Leka)	1.00	0.98	0.94	0.96	0.91	0.83	1.00	1.00	0.98	1.00	1.00	1.00	1.5	6.4	18	24	39	84
AZAM																		
AZAM (B. Lites)	1.00	0.99	0.98	0.96	0.98	0.95	1.00	1.00	0.99	1.00	1.00	1.00	1.5	6.3	17	0.13	24	73

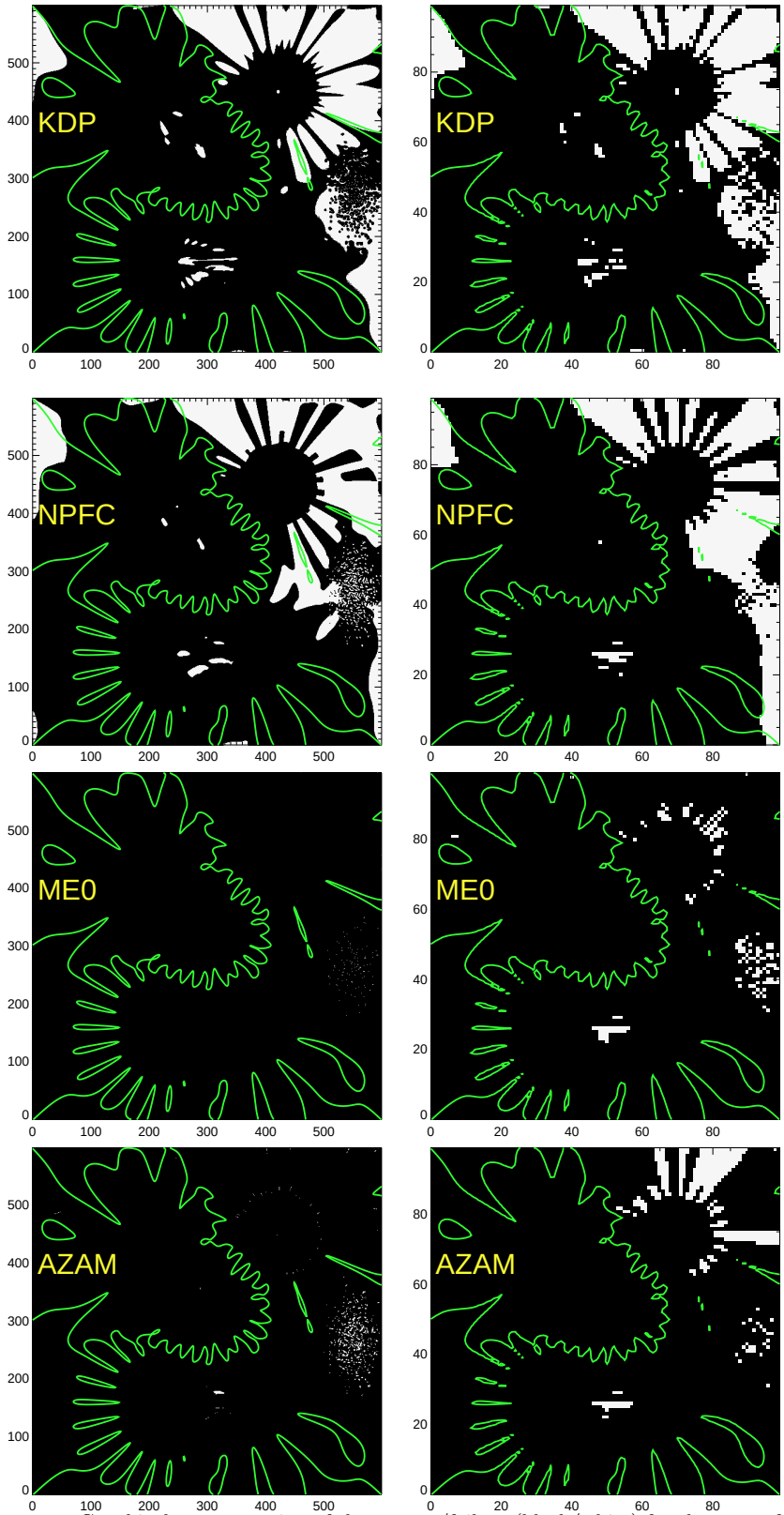

Figure 11. Graphical representation of the success/failure (black/white) for the tests of spatial resolution, for representative algorithms. (Left/right:) the 0.15″ and 0.9″ cases respectively, including a contour of the magnetic neutral line for reference (compare Figure 4). (From top): results for a Potential-Field acute-angle method, the NPFC2 method, a Minimum-Energy method, and the `AZAM` utility (by Kubo and Lites, respectively). While not shown, results from other Potential-Field Acute-Angle methods basically resemble the top figures, the UHIM results most closely resemble those from NPFC2.

Table 4. Results for limited-resolution cases.

	$\mathcal{M}(\mathrm{a,s})_{\mathrm{area}}$			$\mathcal{M}(\mathrm{a,s})_{J_z}$			$\mathcal{M}(\mathrm{a,s})_{B_\perp>100\mathrm{G}}$			$\mathcal{M}(\mathrm{a,s})_{B_\perp>500\mathrm{G}}$			$\mathcal{M}_I$ (10^{13}A)			$\mathcal{M}_{\Delta\mathbf{B}}$		
	0.15″	0.3″	0.9″	0.15″	0.3″	0.9″	0.15″	0.3″	0.9″	0.15″	0.3″	0.9″	0.15″	0.3″	0.9″	0.15″	0.3″	0.9″
"Answer field"													0.34	0.46	0.78	1.8	4.7	20
Potential, FFT																		
YLP (Y. Liu)		0.69	0.69		-1.6	-0.08		0.90	0.90		0.96	0.96		2.5	2.1		123	134
KLP (K. Leka)	0.83	0.83	0.83	-1.7	-1.1	0.12	0.94	0.94	0.94	0.98	0.98	0.98	2.0	2.1	1.8	75	78	90
JJP (J. Jing)		0.80	0.80		-1.2	0.16		0.93	0.93		0.98	0.98		2.2	1.8		82	94
GBP (G. Barnes)	0.83	0.83	0.84	-1.7	-1.1	0.18	0.94	0.94	0.94	0.98	0.98	0.98	2.1	2.1	1.7	77	78	92
Non-potential Field Calculation II																		
NPFC2 (G. Barnes)	0.84	0.85	0.84	-2.2	-0.43	0.42	0.94	0.94	0.94	0.98	0.98	0.98	2.3	1.4	1.4	75	70	82
Pseudo-Current																		
PCM (A. Gary)	0.78	0.80	0.72	-2.2	-2.0	-0.1	0.90	0.91	0.92	0.95	0.94	0.95	2.4	3.0	2.1	116	110	119
UH Iterative																		
UHIM (K. Leka)	0.87	0.87	0.83	-0.65	-0.26	0.51	0.95	0.95	0.95	0.99	0.98	0.99	1.4	1.3	1.3	57	66	79
Minimum Energy																		
ME0 (Crouch&Barnes)	1.0	1.0	0.98	0.95	0.81	0.82	1.00	1.00	0.99	1.00	1.00	1.00	0.34	0.48	0.92	1.8	5.3	27
ME2 (K. Leka)		0.86	0.89		-0.002	0.62		0.94	0.96		0.98	0.99		1.0	1.1		75	81
AZAM																		
AZAM (B. Lites)		0.99	0.96		0.56	0.68		1.00	0.98		1.00	0.99		0.61	1.1		6.9	40
AZAM (M. Kubo)	1.00	0.99		0.21	0.51		1.00	1.00		1.00	1.00		0.77	0.66		3.5	7.2	

We find that there are few obvious trends with spatial resolution, but we note that we present a sparse sampling. We find that the weak trends differ somewhat according to an algorithm's approach: arriving at a solution by way of an acute angle test to a comparison field computed from the boundary, or arriving at a solution by way of simply smoothing the field or globally optimizing a prescribed functional.

For the latter methods, such as ME[0,2] and AZAM, there is a trend evident both spatially and from the metrics of a general decrease in performance with worsening spatial resolution. In contrast, reference-field methods, such as all potential-field acute-angle methods, fail in almost identical areas as spatial resolution is varied. Methods which do not use global optimization, such as NPFC2, PCM, and UHIM, perform similarly with varying spatial resolution based on the metrics, but fail in somewhat different areas as spatial resolution changes.

The interpretation of the metrics requires some care. In some cases, the spatial resolution degradation did not translate to an obvious degradation in the metrics. Those metrics which compute a (weighted) fraction of the total resolved correctly, such as $\mathcal{M}(\mathrm{a,s})_{\mathrm{area}}$ or $\mathcal{M}(\mathrm{a,s})_{B_\perp>100\mathrm{G}}$ are unable to distinguish performance differences between the spatial resolutions. As an example, 3 of the 6 metrics are almost identical between the tests for the potential-field acute-angle cases.

The metrics based on the vertical current, $\mathcal{M}(\mathrm{a,s})_{J_z}$ and $\mathcal{M}_I$, seem to indicate that methods may improve with worsening spatial resolution. For example, recalling that $\mathcal{M}(\mathrm{a,s})_{J_z} = 1.0$ is best, most methods have $\mathcal{M}(\mathrm{a,s})_{J_z} < 0$ for the $0.15'', 0.3''$ cases but $\mathcal{M}(\mathrm{a,s})_{J_z} > 0$ for $0.9''$. However, by visually examining the spatial right/wrong maps (Figure 11) we see that this improvement is not really the case.

Instead, decreasing spatial resolution effectively increases the distance between pixels; the magnitude of the horizontal field may or may not (depending on how rapidly the underlying field is varying) decrease by a similar amount. Thus, current computed by finite differences will *generally* have a smaller magnitude as $\Delta\, x$ increases. In addition, the $\mathcal{M}(\mathrm{a,s})_{J_z}$ metric is sensitive to line-currents: with fewer pixels available between the remnants of the fine-scale structure, there are fewer pixels across which to produce strong line currents, and the metric appears to improve with worse spatial resolution.

The $\mathcal{M}_I$ metric provides mixed messages, but in this case we also have the results for the "answer" fields themselves. While small, the hares show an increase in total current for worsening spatial resolution. `AZAM`, ME0, follow this trend, while the rest all suffer from the misleading trend of decreasing $\mathcal{M}_I$ with worsening spatial resolution while having overall substantially higher $\mathcal{M}_I$ scores. Recalling that the underlying model field is potential, for which $\mathcal{M}_I = 0.0$, this metric provides a good evaluation when fully understood.

The $\mathcal{M}_{\mathbf{\Delta B}}$ metric indicates how much additional "noise" is present in the magnetogram due to incorrect ambiguity resolution. For comparison, we include this metric computed by comparing the hare, or answer field, with the model field, finding that the three spatial resolution cases have 1.8/4.7/20 G of noise respectively imparted in the field solely from the binning and spectral inversion itself. All ambiguity resolution methods add additional noise, and all follow the

trend of higher $\mathcal{M}_{\Delta\mathbf{B}}$ for lower spatial resolution. Some methods impart minimal additional noise (ME0, `AZAM`), while others add substantial noise levels (the potential-field acute-angle methods). This appears to be one of the most sensitive and clearest metrics for interpreting the effects of spatial resolution.

As with the Poisson-noise added cases, the physical assumptions invoked by the algorithms may no longer be consistent with the boundary. As an example, despite the underlying field being a current-free potential field construct, the acute-angle algorithms based on a potential field did not perform particularly well. Algorithms based on minimizing quantities such as the vertical current could be affected as well: as discussed earlier, the intensity-weighted mixing of the polarization spectra sampled on a sparse discrete grid enhances the inter-pixel discontinuities, resulting in significant inferred unsigned vertical current where there is none in the underlying field.

The structure at the top/right of the model, with the ring of azimuth centers, was particularly troublesome; contributor Lites stated, "there is no obvious `AZAM` good solution for the upper right" regarding the $0.9''$ case. The effects of degrading spatial resolution are dramatic here, in the sense that the inverted points deviate noticeably from a directly binned model field: the "azimuth centers" become unrecognizable. In addition, when one computes a reference field (potential or linear force-free) using solely the $B_{\|}$ boundary, the fine-scale structure is completely lost. As a result, a reference field computed in this way behaves as if this should be a single negative-polarity sunspot, with all associated horizontal field heading toward the umbral center. The assumptions used in computing the reference field to supplement the inadequate information provided by just the boundary observation, lead the solutions astray.

Methods which performed a global optimization, or those which relied solely on spatial smoothness without regard to a reference field (such as `AZAM` at the $0.3''$ resolution) were able to recover this somewhat counter-intuitive structure. Spatially, ME[0], `AZAM`, generally performed well on this structure. On the other hand, methods which relied on a reference field for an acute-angle solution systematically failed in the outer parts of the north/west sunspot. Hybrid methods such as UHIM and NPFC2, which include treatment of non-potential fields or iterative minimization of divergence, performed with mid-quality results.

Still, smoothing or global optimization methods were not perfect; *all* methods failed to some degree in the "plage area", even the bin-5, $0.15''$ solution from ME0. This model "plage" is a collection of very small but fully resolved (before binning) azimuth centers with significant horizontal field, randomly distributed subject to constraints on overlap. They have none of the "reverse" field challenges of the north/west structure; this area is purely a test of the effects of spatial averaging. With the underlying field varying rapidly with position, it is the most problematic with regards to the assumption that finite-differences suitably approximate the true derivatives, as seen by the area's large values of $\mathcal{D}$ (see Figure 8). These structures are prone to a varying degree of mixing of the polarization spectra and do not conform to (are not successfully resolved by) any single method's assumptions at any spatial resolution tested here. We demonstrate that while not attempting to provide a fully developed model of solar plage areas, the "plage" area reproduces the salient features of spectropolarimetric observations.

One must wonder, given that all methods had difficulty on this basically simple magnetic field construct, how to best interpret real observations of unresolved structure.

5. Conclusions

We subjected a variety of algorithms for the ambiguity resolution in vector magnetic field data to systematic challenges, and evaluated their performance in the presence of observational noise. Two different noise sources were tested: photon noise and unresolved structure. All observational data are susceptible to these two types of noise. Even at very high spectropolarimetric precision and very high spatial resolution, there *will* be Poisson-distributed photon shot-noise, and there *will* be unresolved structures in each resolution element. The question is one of the level of noise and the extent to which an algorithm's performance degrades.

As in Paper I, we find that even acute-angle potential-field methods can give different results according to the details of the potential field calculation and the treatment of the side boundaries. None of the acute-angle potential-field algorithms performed substantially better or worse with increasing noise (whether photon or spatial-resolution) than expected. As in Paper I, the acute-angle potential field methods fared moderately, and can be considered a standard against which more sophisticated models should be compared.

We find that almost all algorithms perform less well with an increased level of photon noise, as expected. The different algorithms have different sensitivity, and one in fact performs better at higher noise levels. No method appears to widely propagate a bad solution from a certain noisy pixel, and no method completely succeeds in the face of the (unrealistic) challenge of determining the azimuthal direction when the signal-to-noise ratio for the transverse field is less than unity.

Indeed, it is in the small-magnitude areas of the transverse field where photon noise is most troublesome, and these tests remind us that such areas are *not* limited to plage or quiet-sun areas, but can occur mid-sunspot as well. In these areas, the noise appears as random vertical current which can have different effects depending on whether it occurs in weak-field areas or in the middle of a sunspot.

For all ambiguity-resolution algorithms, the existence of photon noise simply means that there is inferred current where the underlying field has none, and there is limited information to be used for the ambiguity-resolution. This implies that the assumptions can be invalid when the polarimetric signal is dominated by noise, even when the assumptions are valid for the true field. Global minimization of a physics-based function was not a guaranteed success. Those algorithms which relied upon a spatial smoothing in their approach appeared to perform the best in areas where photon-noise dominates.

Isolating the effects of unresolved structure produced somewhat parallel results. Finite spatial resolution spatially mixes the polarization spectra in an intensity-weighted non-linear manner. When this mix is inverted, unresolved structures are not necessarily "smoothed out"; the inversion itself will introduce noise and the increased pixel-spacing introduces large apparent inter-pixel

jumps. Significant vertical current is then inferred where there was none at full resolution, and the "observed magnetogram" may no longer be consistent with $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$. Does this constitute an inappropriate test, then? Not in the least. In this case we know what the changes are from a constructed field for which $\boldsymbol{\nabla} \times \mathbf{B} = 0$ and $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$, and thus can test how sensitive the algorithms are to approximating derivatives from discretely-sampled data. For solar observations, we *assume* that the resulting magnetic field map, inverted from spectropolarimetric data at finite resolution, is a valid field. We show here that it may not be so.

A result from Paper I is confirmed: all methods employ assumptions in light of the insufficient information provided by a single-height observation. If the assumptions from which a reference field is constructed do not match the true field, the results are usually not good. Details of the area outside the field of view (horizontally or vertically) and its effects on the structure within the field of view can be problematic.

Moreover, different implementations of the same basic assumption can lead to different results. ME0 and NPFC2 use the same basic assumption of deriving $\partial B_z / \partial z$ by the use of a potential-field extrapolation, but the subsequent use of the derived information is different, as are the results. The same can be said for the potential-field acute-angle methods: the same basic assumption is used, but the treatment of the side boundaries in particular can lead to very different results.

All methods performed very well in areas of strong transverse fields, with $\mathcal{M}(\mathrm{a},\mathrm{s})_{B_\perp > 500\mathrm{G}} \geq 0.95$. No method ever produced a perfect solution to any of the cases tested when noise was present; no method ever perfectly resolved the ambiguity in areas which were not spatially resolved. The new "average vector difference" metric demonstrated that spatial resolution effectively adds noise to a magnetogram. We find that low-scoring algorithms worsen this effect, while algorithms based on global minimization of $\boldsymbol{\nabla} \cdot \mathbf{B}$ and J_z perform best against all of these challenges.

In summary, in light of observational sources of noise, the potential-field acute-angle methods once again were the "'standard" against which more sophisticated methods were judged. Hybrid methods such as UHIM and NPFC2 performed with good but not great results. Algorithms which performed a global optimization of a physics-based functional overall performed best of the automated methods. Results for interactive methods such as `AZAM` were dependent on the user but generally very good (when available).

The behavior of the algorithms in light of observational noise sources points to an observational approach which can be taken with today's instruments: first invert polarization spectra which are obtained on the finest spatial scale possible, then perform the ambiguity resolution. If spatial binning is considered advantageous, there may be cases where the binning should be performed *not* on the spectra, but on the final ambiguity-resolved vector field map. Subject to appropriate signal/noise limits, with this approach the noise added due to spectral mixing is be minimized, errors in the inversion and ambiguity-resolution affect the smallest observed area possible, and necessary assumptions are the least challenged by the nature of discrete data. This is contrary to conventional

wisdom of binning spectra prior to a spectral inversion, but might be taken under advisement in certain specific signal/noise scenarios. Given the results of the tests performed herein, we put forth this guarded suggestion.

Acknowledgements We thank Ms. Janet Biggs for logistical support during the workshop held at NWRA/CoRA Division, and the community and our agency program officers for their support and programmatic flexibility after the loss of Dr. Metcalf. We thank H. N. Wang and Y. J. Moon for submitting solutions for the photon-noise case for analysis, B. W. Lites and M. Kubo for solutions to cases using `AZAM`, and J. Li and M. K. Georgoulis for enlightening conversations. KDL, GB and AC acknowledge funding from NASA/LWS under contracts NNH05CC49C and NNH05CC75C and the NWRA subcontract from SAO under NASA NNM07AB07C. The results for the ME0 algorithm derives in part from codes tested and developed with funding to NWRA through the HMI project at Stanford University, NASA Grant NAS5-02139.

References

Auer, L.H., Heasley, J.N., House, L.L.: 1977, *Solar Phys.* **55**, 47.
Borrero, J.M., Solanki, S.K.: 2008, *Astrophys. J.* **687**, 668.
Canfield, R.C., de La Beaujardière, J.F., Fan, Y., Leka, K.D., McClymont, A.N., Metcalf, T., Mickey, D.L., Wülser, J.P., Lites, B.W.: 1993, *Astrophys. J.* **411**, 362.
Crouch A.D., Barnes G., Leka K.D.: 2009, *Solar Phys.* **submitted**.
Georgoulis, M.K.: 2005, *Astrophys. J. Letters* **629**, L69.
Ichimoto, K., Lites, B., Elmore, D., Suematsu, Y., Tsuneta, S., Katsukawa, Y., Shimizu, T., Shine, R., Tarbell, T., Title, A., Kiyohara, J., Shinoda, K., Card, G., Lecinski, A., Streander, K., Nakagiri, M., Miyashita, M., Noguchi, M., Hoffmann, C., Cruz, T.: 2008, *Solar Phys.* **249**, 233.
Keller C.U., The Solis Team : 2001, In: Sigwarth, M. (ed.) *Advanced Solar Polarimetry – Theory, Observation, and Instrumentation*, *ASP Conf. Ser.*, **236**, 16.
Labonte, B.: 2004, *Solar Phys.* **221**, 191.
Labonte, B., Mickey, D.L., Leka, K.D.: 1999, *Solar Phys.* **189**, 1.
Leka K.D., Barnes G., Crouch A.D.: 2009, In: *Second Hinode Science Meeting. ASP Conf. Series.*
Lites B.W., Elmore D.F., Streander K.V., Sankarasubramanian K., Rimmele T.R., Sigwarth M.: 2003, In: Trujillo-Bueno, J., Sanchez Almeida, J. (eds.) *Solar Polarization 3*, *ASP Conf. Ser.*, **307**, 324.
Metcalf, T.R.: 1994, *Solar Phys.* **155**, 235.
Metcalf, T.R., Leka, K.D., Barnes, G., Lites, B.W., Georgoulis, M.K., Pevtsov, A.A., Gary, G.A., J ing, J., Balasubramaniam, K.S., Li, J., Liu, Y., Wang, H..N., Abramenko, V., Yurchyshyn, V., Moon, Y.J.: 2006, *Solar Phys.* **237**, 267.
Mickey, D.L., Canfield, R.C., LaBonte, B.J., Leka, K.D., Waterson, M.F., Weber, H.M.: 1996, *Solar Phys.* **168**, 229.
Parker, E.N.: 1996, *Astrophys. J.* **471**, 485.
Sakurai, T., Ichimoto, K., Shinoda, Y., Noguchi, M., Hiei, E., Li, T., He, F., Mao, W., Lu, H., Ai, G., Zhao, Z., Kawakami, S., C., C.J.: 1995, *Publ. Astron. Soc. Japan* **47**, 81.
Sankarasubramanian K., Lites B., Gullixson C., Elmore D., Hegwer S., Streander K., Rimmele T., Fletcher S., Gregory S., Sigwarth M.: 2006, In: Casini, R., Lites, B.W. (eds.) *Solar Polarization 4*, *ASP Conf. Ser.*, **358**, 201.
Scherrer P.H., Hoeksema J.T., The HMI Team : 2006, 36th COSPAR Scientific Assembly, Abstract No. 1469 (in CDROM)..
Skumanich, A., Lites, B.W.: 1987, *Astrophys. J.* **322**, 473.

Stenflo, J.O., Keller, C.U.: 1997, *Astron. Astrophys.* **321**, 927.
Titov, V.S., Priest, E.R., Demoulin, P.: 1993, *Astron. Astrophys.* **276**, 564.
Tsuneta, S., Ichimoto, K., Katsukawa, Y., Nagata, S., Otsubo, M., Shimizu, T., Suematsu, Y., Nakagiri, M., Noguchi, M., Tarbell, T., Title, A., Shine, R., Rosenberg, W., Hoffmann, C., Jurcevich, B., Kushner, G., Levay, M., Lites, B., Elmore, D., Matsushita, T., Kawaguchi, N., Saito, H., Mikami, I., Hill, L.D., Owens, J.K.: 2008, *Solar Phys.* **249**, 167.
Venkatakrishnan, P., Gary, G.A.: 1989, *Solar Phys.* **120**, 235.