1. INTRODUCTION

Clusters of galaxies are excellent cosmological laboratories (Allen et al. 2008; King 2007; Mannucci et al. 2008). For example, the mass function of clusters is a sensitive probe of cosmological parameters like $\Omega_m$ and $\sigma_8$ (Rines et al. 2007), and its observed evolution is an important test of theories of structure formation (Gunn & Gott 1972; Giocoli et al. 2007; Horellou & Berge 2005; Cooray & Sheth 2002). The geometrical shape of cluster dark matter halos provides valuable information on the intracluster gas distribution (Flores et al. 2005, 2007). While simulations predict the central density distribution of matter in clusters to follow a Navarro-Frenk-White (NFW) profile, it is debatable whether observations suggest that clusters have a central core (Sand et al. 2004; Voigt & Fabian 2006).

ACDM structure formation theories also predict that massive dark matter halos assemble from the hierarchical merging of lower mass subhalos. As noted by several authors (Moore et al. 1999; Klypin et al. 1999), the number of subhalos that survive in N-body simulations is much greater than the number of dwarf galaxies observed in the Milky Way and the Andromeda galaxy. On cluster scales, such discrepancies are not observed. Thus, the subhalo mass function in clusters is an important probe of the CDM theory in this mass scale.

High-resolution, accurate measurements of cluster mass maps are thus highly desirable. Gravitational lensing is a powerful tool to probe the projected mass map of the clusters independent of the internal dynamics, and has already been widely applied to mapping mass distribution in clusters (Wittman et al. 2001; Hoekstra et al. 2001; Gray et al. 2002; Taylor et al. 2004; Broadhurst et al. 2005b; Leonard et al. 2007; Okura et al. 2008; Heymans et al. 2008). Some researchers (Natarajan & Springel 2004; Natarajan et al. 2007b) have used the individual galaxy-galaxy lensing signal to estimate individual galaxy masses and, thus, produce a parametric mass reconstruction of the cluster. Others have used the weak signal to characterize the overall potential from the cluster without recourse to parametric models (Wilson et al. 1996; Hoekstra et al. 1998, 2004; Natarajan & Refregier 2000).

Given the importance of accurately measuring the mass, shape, and substructure of individual clusters and given the enormous expense of long–time exposure observations of clusters, it is extremely important to maximize the signal-to-noise ratio from a particular data set and to produce high-resolution maps of substructure within individual clusters. Current mass reconstruction techniques are ill equipped to handle multiscale data sets or clusters with significant clumpiness or cuspliness, or are jury-rigged to do so. In this paper we propose a shear-only reconstruction for reconstructing galaxy clusters and identify some strengths and limitations of this approach.

Our outline is as follows. In § 2 we give a brief review of the essential lensing formalism and lay out our notation for the rest of the work. In § 3 we describe our method (grid-based) techniques for reconstructing galaxy clusters and identify some strengths and complications. In § 4 we propose PBL. We then apply this new method to simple simulated clusters of single- and double-peak softened isothermal spheres and the “Bullet Cluster” (1E 0657–56) in § 5. We conclude in § 6 with a discussion of future prospects, including how additional strong-lensing and flexion information can be incorporated into PBL.

2. BACKGROUND

Before delving into technical details of our method we would like to introduce the basic lensing notation to be used throughout the paper. Following Bartelmann & Schneider (2001), we consider a surface mass density $\Sigma(\theta)$, where $\theta$ is the angular position in the lens plane. Convergence or dimensionless surface mass density is defined as

$$\kappa(\theta) = \frac{\Sigma(\theta)}{\Sigma_{cr}},$$

(1)
where

\[ \Sigma_{\text{cr}} = \frac{c^2}{4\pi G} \frac{D_s}{D_d D_{ds}}, \tag{2} \]

\( D_s, D_d, \) and \( D_{ds} \) are the angular diameter distances between the observer and the source, the observer and the lens, and the lens and the source, respectively. For convenience, we define a fiducial critical density for a source plane at \( z_s = z_{ds} = \infty \), and all models are scaled to this standard.

The convergence is related via a Poisson-like equation to a normalized potential,

\[ \nabla^2 \psi(\theta) = 2\kappa(\theta). \tag{3} \]

Here and throughout this paper, all derivatives are in angular units in the lensing plane. A single light beam is deflected by

\[ \alpha = \nabla \psi. \tag{4} \]

The lens equation relates the source position \( \beta \) to the image position(s), \( \theta \), as

\[ \beta = \theta - \nabla \psi. \tag{5} \]

When the lensing potential does not vary appreciably across the source, the lens mapping can be linearized. The transformation between the source and the image is given by the Jacobian matrix

\[ \mathbf{A}(\theta) \equiv \frac{\partial \beta}{\partial \theta} = (\delta_{ij} - \psi_{ij}) = \begin{pmatrix} 1 - \kappa & -\gamma_2 \\ -\gamma_2 & 1 - \kappa \end{pmatrix}. \tag{6} \]

From this, we see that distortions in shape are well described in terms of shear, which is related to the lensing potentials through the relations

\[ \gamma_1 = \frac{1}{2} (\psi_{11} - \psi_{22}), \tag{7} \]

\[ \gamma_2 = \psi_{12}, \tag{8} \]

using Einstein convention for derivatives.

The radial eigenvalue is given by \( \lambda_+ = 1 - \kappa + |\gamma| \), and the tangential eigenvalue is given by \( \lambda_- = 1 - \kappa - |\gamma| \). The matrix is singular where \( \lambda_\pm = 0 \). These points define the critical curves of the lens.

Third-order corrections to the lensing potential become non-negligible when the lensing potential varies across the image. These observables, the gravitational flexion, were derived in Goldberg & Bacon (2005). They are more colloquially referred to as the “bananeness” (Schneider & Er 2008) or bending of an image. It is based on the third angular derivatives of the potential (Bacon et al. 2006) given by

\[ \mathcal{F} = (\gamma_{11} + \gamma_{22}) + i(\gamma_{21} - \gamma_{12}) \tag{9} \]

\[ \mathcal{G} = (\gamma_{11} - \gamma_{22}) + i(\gamma_{21} + \gamma_{12}). \tag{10} \]

Each of the deflection (\( \alpha \)), the shear (\( \gamma_1, \gamma_2 \)), the convergence (\( \kappa \)), and the flexion (\( \mathcal{F}, \mathcal{G} \)) are linear functions of the potential field. While the discussion in this paper primarily focuses on measurements with sources at fixed, infinite redshift, it should be noted that each of these terms scales as

\[ \kappa(z_s) = Z(z_d, z_s)\kappa(z_s = \infty), \tag{11} \]

where

\[ Z(z_s) = \frac{D_{ds}}{D_s}. \tag{12} \]

3. GRID-BASED CLUSTER LENSING

Mass reconstruction studies have been very successful on cluster scales (Bartelmann & Schneider 2001; Clowe & Schneider 2002; Hoekstra et al. 2002; Broadhurst et al. 2005a; Okura et al. 2008 and references therein). Because these systems typically contain many lensed images, the shear signal can be extracted with high significance. In this section we describe an important class of cluster inversion techniques which reproduce the convergence field on a grid. Our specific choices of grid-based techniques include those which have already been extended to include strong-lensing information with nonparametric models and, thus, provide a fertile basis for comparison. Further, there are many variants even within the subcategory of grid-based reconstruction techniques. We focus primarily on their commonalities, as exemplified by those discussed in Bradač et al. (2005a, 2005b) and Cacciato et al. (2006). We focus on methods in which various scalar fields \( \{\psi, \kappa\} \) are defined on a Cartesian grid and minimized according to the criteria described below. In so doing, we note some interesting exceptions: Diego et al. (2005) and Saha et al. (2001), who describe an adaptive mesh technique for refining the field on different resolution scales, and Marshall et al. (2002) and Marshall (2006), who use a variable smoothing scale for their weak-lensing mass reconstruction.

3.1. Weak Lensing on Grids

The standard approach to lensing arclet inversion (Luppino et al. 1999) has been to measure the ellipticity of observed images as an unbiased estimator of the reduced shear,

\[ \langle \epsilon \rangle = g \equiv \frac{\gamma}{1 - \kappa}. \tag{13} \]

For relatively weak fields (\( \kappa \ll 1 \)), this is very nearly a direct estimate of the shear and can perform a direct finite inversion to estimate the density field.

In recent years, there has been a flurry of work on optimal methods for nonparametric cluster mass reconstructions (Bradač et al. 2005a, 2005b; Natarajan & Springel 2004). In general, these papers focus on estimating the potential \( \{\psi\} \) or convergence \( \{\kappa\} \) fields of a cluster by a \( \chi^2 \) minimization analysis. Both the shear and the convergence are linear functions of the potential field. Thus, if a model potential field, \( \{\psi\} \), is defined on a grid, then the shear at some grid cell, \( i \), may be expressed as a linear combination of potential,

\[ \gamma_{1i} = G_{ij}^{(1)} \psi_j, \tag{14} \]

with a similar expression for the convergence, \( \kappa \), and the imaginary component of the shear, \( \gamma_2 \). We refer to these below simply as \( \gamma_i(\{\psi\}) \), since we wish to remind the reader that the estimate of the shear is an explicit function of the test potential field. Because these fields are combinations of second derivatives of the potential field, the \( G^{(1)} \) matrix and the others are easy to compute.
using finite differencing and are extremely local. A very good graphical representation of the finite-difference operators can be found in Bradač et al. (2005b).

In the weak-field limit, the complex ellipticity of a lensed galaxy is a linear, albeit noisy, estimator of the complex shear field. The principle component of noise is the intrinsic ellipticities of galaxies which follow a Gaussian distribution with a standard deviation of $\sigma_e \approx 0.3$ for each component. The large variance in intrinsic ellipticities necessitates averaging over many images so that the noise in a single grid cell is zero or applying an artificial smoothing scale to a more finely gridded mesh. For a weak-lensing–only calculation, a $\chi^2$ minimization is performed on

$$\chi^2_W = \sum_i \left[ \frac{\gamma_i(\{\psi\})}{1 - \kappa_i(\{\psi\}^{(a-1)})} - \epsilon_i \right]^2,$$

where the estimate of $\kappa$ is taken from the previous iteration of the potential field, and thus, the model rapidly converges to a maximum-likelihood solution to the potential field.

### 3.2. Strong Lensing

A number of researchers, including Bradač et al. (2005a, 2005b), have noted that a similar grid-based formalism may be used with strong-lensing signals. Strong+weak (S+W) reconstructions using both shear fields and the positions of multiply imaged sources can be used to accurately reconstruct both the cores and halos of clusters. While our current PBL implementation, described in § 4, does not currently incorporate strong-lensing analysis, we introduce this component of grid-based lensing reconstructions to illustrate how directly a strong-lensing analysis could be incorporated into PBL.

Strong lensing by clusters produces an especially elegant result, because if, say, two images are observed at positions $\theta^A$ and $\theta^B$, then it must be true that both images originated at the same (unknown) position in the sky. Thus, we have a simple relation,

$$\theta^A - \alpha(\theta^A) = \theta^B - \alpha(\theta^B).$$

The appeal of this relationship is that it is fundamentally linear, and thus, the angular separation between the two images (itself, a measurable quantity) can be directly related to the difference in the first derivatives of the potential at two different points in the field.

As above, the local derivatives can be computed as

$$\alpha_{ij} = A_{ij}^v \psi_j,$$

with a similar expression for the $y$-component of the displacement. The matrix elements of $A$ are easy to compute, as they are simply the first derivative in a simple grid-based second-order difference scheme. More generally, we can express this as $\alpha_{ij}(\{\psi\})$. Thus, an additional $\chi^2$ term can be added,

$$\chi^2_s = \sum_{i, \text{pairs}} \left[ \frac{[\alpha^x(\{\psi\}) - \alpha^y(\{\psi\})] - (\theta^A - \theta^B)}{\sigma_i^2} \right]^2,$$

and minimized either independently or simultaneously with the weak-lensing component.

### 3.3. Regularization

Using a $\chi^2$ minimization technique discussed in § 3.1, it is possible to get a checkerboard pattern due to independent noise in the two components of ellipticity. This requires the addition of a regularization term to $\chi^2$ to suppress this noise.

Scale refinement is also necessary in cases in which S+W lensing signals are combined. To make this argument concrete, consider a toy isothermal sphere model of a cluster with a one-dimensional velocity dispersion of $600 \text{ km s}^{-1}$. Each multiply imaged pair will be separated by twice the Einstein radius, about $20''$ in this case. This represents the minimum necessary resolution in the reconstruction to say anything about strong lensing.

On the other hand, even very efficient space-based weak-lensing analyses of clusters seldom yield more than approximately 100 images arcmin$^{-2}$. Using a simple Poisson noise estimate, we may achieve uncertainty of $\sigma_e = 0.06$ only with images binned on scales larger than $30''$ on a side. Smaller binnings will naturally yield larger uncertainties. The simple grid-based method cannot both capture the weak-lensing signal to high accuracy and resolve the strong-lensing regime. In order to deal with this issue, different investigators have used different regularization techniques.

One method is to use a series of finer and finer griddings, and at each successive level of refinement, the convergence field from the previous level is matched as closely as possible. The Bradač et al. (2005b) S+W technique uses this method, with the weighting parameter selected to provide a $\chi^2$ per degree of freedom equal to 1, such that

$$R = \eta \sum_i \left[ (\kappa_i^{(a)} - \kappa_i^{(a-1)})^2 \right],$$

where $\kappa_i^{(a-1)}$ represents the estimated convergence on the previous, coarser gridding, and where $\kappa_i^{(0)} = 0$. We use this form explicitly in § 5 where we test the PBL method and contrast it with grid-based reconstruction methods.

### 3.4. Some Questions

Grid-based reconstructions have produced some excellent measurements; however, there remain a number of complications. First, grid-based techniques are really optimized to measure a single scale, the grid spacing. However, as we discuss above, in many interesting systems, both the structure and information are hierarchical. An optimal technique should provide higher resolution in regions of greater information content.

Moreover, the smoothing and weighting of the strong lensing, weak lensing, and regularization are created on an ad hoc basis. The ideal smoothing scale should be variable and such that the signal-to-noise ratio of the reconstructed field is similar in every smoothed cell.

Third, the information from the image ellipticity can only be inverted outside the critical curves of the lenses. Inside the (tangential) critical curve (Schneider et al. 1992; Petters et al. 2001; Schneider & Weiss 1992; Hoekstra et al. 2004) there is an abrupt switch in parity of the induced ellipticity of an image. More plainly, in the regime $|\gamma| > 1 - \kappa$, the ellipticity is related to the shear via

$$\langle \epsilon \rangle_{\text{strong}} = \frac{1}{g}.$$

As discussed in § 4.4, this produces a discontinuity in the ellipticity as a function of $\kappa$ and $\gamma$. No simple linear minimization scheme, even an iterative one, will converge to the “strong lens” solution if one starts with a “weak lens” initial guess for the local potential field.
4. PARTICLE BASED LENSING

In this section we introduce a new technique called Particle Based Lensing (PBL), which has the ability to combine the disparate lensing scales in a coherent way without requiring a regularization scheme. Several of the concerns discussed above have to do with the method of discretizing the data for the reconstruction of the lens potential. In order to address this, we turn to a technique which is widely used in another area of astrophysics in which information must be analyzed on a wide range of physical scales—numerical N-body simulations. Smoothed particle hydrodynamics (see, e.g., Monaghan 2005 for a recent review) is used in the modeling of a wide range of physical systems including planets (Woolfson 2007), star formation (Springel & Hernquist 2003; Nagamine et al. 2004), and galaxy formation (Kaufmann et al. 2007). The mathematical details of PBL can be complicated; hence, we have made our codes for the method public1 through our Web site.

Before getting into the details, however, it is important to emphasize what PBL is and is not. PBL is a new way of discretizing and describing a reconstructed field. Moreover, it includes a metric for comparing a reconstructed model to the observed data. Everything we describe below is aimed at demonstrating why this model and metric are ideal for lensing systems with uneven information content. While the current code excludes a metric for comparing a reconstructed model to the background images and type of constraint (e.g., ellipticity, position), the worked examples are based on weak-lensing data only, which is widely used in another area of astrophysics in which information content is the relative offset of the test point from galaxy n. In order to make the field as continuous as possible, we may expand the local potential field around the position of any lensed image (ψn, in this case) to arbitrary order,

where θ is the relative offset of the test point from galaxy n.

As with grid-based lensing, the local derivatives are composed of a linear combination of the potentials at each grid point. That is,

ψn,j = D(n)jψn,m,

ψn,jk = D(n,k)ψn,m,

and so on for arbitrarily higher derivatives. In reality, we typically extend the D(ν) matrices up to third order, where ν corresponds to two matrices for first derivatives (displacement field), three for second derivatives (shear and convergence), and four for third derivatives (flexion). Here we use the Einstein summation convention, the sum over m runs from 1 to Ng.

In terms of the D(ν) matrices, equation (20) may be rewritten as

ψ(θm) = ψn + ∑ D(ν)nmX(ν)nmψl, (23)

where we are explicitly estimating the potential at the nth galaxy from the local derivatives defined at the nth. We also compactify equation (20) by defining

X(1)nm = θxx - θxx,

X(2)nm = θxy - θxy,

X(3)nm = (θxx - θxx)2, (26)

and so on.

In order to estimate the derivatives of the potential field near each galaxy, we need to first compute the D(ν) matrices. Since this problem is underconstrained, we solve for these matrices via a χ2 minimization,

χ2 = ∑ W(ψm - ψn - ∑ D(ν)nmX(ν)nmψl)2, (27)

where Wnm is a window function, guaranteeing that only neighboring galaxies affect the potentials of one another. We use a window function of the form

Wnm = W(ψn - θm)/h, (28)

where h is inversely proportional to the signal-to-noise ratio at the nth image position. The smoothing scale can also be chosen to be of the form hnm, i.e., symmetric between the points n and m.

The signal-to-noise ratio is a function of the local density of background images and type of constraint (e.g., ellipticity, positions of multiple images). A similar approach of using a signal-to-noise ratio–dependent smoothing scale has been used in image analysis of X-ray data (Ebeling et al. 2006). In regions where there is a high density of information, the smoothing scale h may be set much lower than in regions of low information density.

This function must be minimized for every matrix element such that

∂2X nm /∂D(ν)nm = 2ψl ∑ X(ν)nmWnm (ψm - ψn - ∑ D(ν)nmX(ν)nmψl) = 0, (29)

But since equation (29) is underconstrained, we may also say

∂2X nm /∂D(ν)nm = 0, (30)

yielding

X(ν)nmX(ν)nmWnmD(ν)nm = X(ν)nmWnm. (31)

1 See http://www.physics.drexel.edu/~dch/PBL.htm.
for all $n, m,$ and $\nu$. This can be solved with a simple matrix inversion, yielding the desired elements for $D^{(\nu)}$. Of course, since the elements are a function only of the positions and weightings of the galaxy images, these elements need only be computed once. The method potentially incorporates higher order derivatives of the potential; thus, combination of strong, weak, and flexion information becomes a relatively straightforward minimization problem.

### 4.2. PBL versus Regularization

One of the major advantages of PBL is that we no longer need to introduce an explicit regularization in order to resolve multi-scale structure in a reconstruction. The various regularization schemes discussed in § 3.3 are not motivated from the associated observations, but are rather derived from assumptions about the mass profile of a cluster motivated by theory and simulations.

However, one of the motivations behind using gravitational lensing is to be able to measure the projected mass without making any assumptions about the physical state of the system. The advantage of using PBL is that we do not need to make any assumptions that go into choosing the regularization term. The smoothing scale of a “pebble” is controlled by $h_n$, which is determined by the local signal-to-noise ratio. This means that the position representing weak-lensing measurements will have a low signal-to-noise ratio and correspondingly a high $h_n$. This is similar to the typical weak-lensing measurement which is done by averaging over a bin size larger than $\sim 30''$. In the case of strong lensing we know the positions of the multiple images for certain, implying a high signal-to-noise ratio and correspondingly a low $h_n$. This can be a few arcseconds, which is the scale at which the strong-lensing structure can be resolved from multiple images. Thus, scales of a few arcseconds can be combined with scales greater than $\sim 30''$ without making any assumptions about the mass profile, rather by taking input from the data.

### 4.3. Estimation of the Potential Field

As with grid-based lensing analysis, in PBL, we use a $\chi^2$ minimization to estimate a maximum-likelihood potential field. In this case, however, we sample the potential at every point and use the local derivatives of the potentials as defined in equation (22) to minimize

$$
\chi^2 = \sum_{i,m} \left[ \frac{\gamma_n^{(i)}(\{\psi\})}{1 - \kappa_m(\{\psi\})} - \xi_n^{(i)} \right]^2 \frac{1}{\sigma_n^2},
$$

where $i$ ranges from 1 to 2 and indicates the real or imaginary components of the shear, reduced shear, or ellipticity. We shall henceforth refer to the first term in the square brackets as $g_n^{(i)}(\{\psi\})$, the estimate of the reduced shear of a model, and the weighting term outside the square brackets as $\omega_n$, yielding

$$
\chi^2 = \sum_{i,m} \left[ g_n^{(i)}(\{\psi\}) - \xi_n^{(i)} \right]^2 \omega_n,
$$

which is the form we refer to from now on.

This is a weak-lensing—only expression. Replacing $g_n^{(i)}(\{\psi\})$ with $1/g_n^{(i)}(\{\psi\})$ gives the strong-lensing counterpart of equation (33). In § 4.4 we discuss how we include this strong-lensing version of the equation.

### 4.4. Interpolated Ellipticities

Linear inversion techniques require that the function to be minimized is smoothly varying over the domain of interest. The ellipticities are given by two functions in the weak- and strong-lensing regimes by equations (13) and (19), respectively. The boundary of the two regimes defines the critical curves where $|g| = 1$, making ellipticities continuous but not differentiable.

The transition between the two regimes can be facilitated if the sources are distributed in redshift, but minimization functions will be much easier if we allow a smoothing of the discontinuities. This is a two-step process, first we need to write equations (13) and (19) in terms of a step function,

$$
\tilde{\xi} \simeq 1 - \mathcal{H}(g) \frac{g}{g^*},
$$

where the function $\mathcal{H}(g)$ is a step function at $g = 1$. We may replace the step function by an approximate smooth function. We define

$$
u = \eta_0 \left( g^2 - \frac{1}{g^2} \right).
$$

Here $\eta_0$ is the free parameter that controls the accuracy of the step function. A higher value of $\eta_0$ makes the step function more accurate. The step function is approximated as (Fig. 1)

$$
\mathcal{H}(u) = \frac{1}{1 + e^{-2u}}.
$$

This approximation replaces the ellipticities only in the neighborhood of the critical curves (discontinuity) by a continuously differentiable function. The problem can now be solved by standard minimization techniques. The interpolated ellipticity function is shown by a dotted line in the right panel of Fig. 1, showing the derivative discontinuity explicitly.

### 4.5. $\chi^2$ Minimization

When we first introduced PBL above, we remarked that it was primarily a way of describing a lens reconstruction in such a way that a small $\chi^2$ would necessarily correspond to a good representation of the underlying field. In practical terms, however, for a reconstruction code to be useful, we need to describe a means of minimizing (or nearly minimizing) the $\chi^2$. Below, we describe our pipeline for fast convergence of a maximum-likelihood solution.

While PBL is a nonparametric reconstruction scheme, it has the useful property that we may start a minimization with any assumed model we like. However, no extra weight is given to our a priori assumptions. At the end of a minimization we may simply use the standard techniques to estimate the likelihood of a particular value of $\chi^2$.

That said, even with the caveat above regarding the smoothing of critical curves, it is very difficult to smoothly vary a solution such that strongly lensed regions are produced. As pointed out by Bradacˇ et al. (2006), a $\chi^2$ minimization process does not ensure reaching a global minimum.

To that end, our initial configuration of $\{\psi\}$ is generated by laying down a small number of singular isothermal spheres (SISs). Since there are a low number of parameters (three for each model sphere), a global minimum may be reached through a combination of trial and error, simulated annealing, or even (for small numbers of spheres), finite sampling. Indeed, one may even use an interpolation of a reconstruction recommended by a grid-based solution. For systems with strong lenses, one may apply the reconstructed field generated by “LensPerfect” (Coe et al. 2008), for example, as a starting point.

We hasten to remind the reader that while this technique will produce the optimum parametric fit, it will not, in general, produce the overall best fit. As a result, further iteration is required.
We have found that by starting with an initial model with well-identified strong-lensing regions, convergence to $\chi^2/\text{DOF} \approx 1$ may be achieved relatively quickly, even if the strong-lensing regions are only approximate. For the current implementation of our code, we use Newton’s method to reach a local minimum. We have found satisfactory, fast convergence for several thousand background sources.

5. TEST APPLICATIONS

In this section we apply PBL to three systems as a proof of concept. In the first, we model a softened isothermal sphere and examine the relative abilities of PBL and grid-based inversion to reconstruct a relatively peaked core. In the second, we model a superposition of two softened isothermal spheres at a given separation as a simple model of a system with substructure. Finally, we reconstruct the “Bullet Cluster” (1E 0657−56; Markevitch et al. 2002, 2004; Clowe et al. 2004, 2006; Bradac et al. 2006), an observed multipeak system of considerable interest. We show that using weak lensing alone, we are able to reconstruct both dark matter peaks.

5.1. Simulation: Softened Isothermal Sphere

5.1.1. Model

We begin by generating a softened isothermal sphere with a potential,

$$\psi = \theta_E \sqrt{\theta^2 + \theta_c^2},$$  \hspace{1cm} (37)

and convergence,

$$\kappa = \theta_E \frac{(\theta^2 + 2\theta_c^2)}{(\theta^2 + \theta_c^2)^{3/2}},$$  \hspace{1cm} (38)

where $\theta_E$ is the Einstein deflection angle given by $4\pi(\sigma/c)^2 D_{ds}/D_s$.

The data are simulated on a unit square field of view. For simplicity we have assumed all sources to be at $z = 1$, with $\theta_E = 0.2$ and $\theta_c = 0.08$. We lens 607 background galaxies and apply an intrinsic ellipticity (noise) with $\theta_{\text{es}} = 0.1$ in each of the principle directions. For all further calculations we use a $\Lambda$CDM cosmology with $\Omega_m = 0.27$ and $\Omega_{\Lambda} = 0.73$. This configuration represents a galaxy cluster at a redshift of $z_{\text{lens}} = 0.4$ with a velocity dispersion of $\sigma_v = 850$ km s$^{-1}$. The field of view is 105' and 0.5 Mpc.

5.1.2. PBL and Grid-based Reconstructions

For the single-peak and the double-peak simulations (see below), we perform both a grid-based reconstruction as well as PBL. We use the regularization suggested by Bradac et al. (2005b) and described in detail in $\S$ 3.3 for the grid-based method. In the case of the single peak, the reconstruction is initially performed on a coarse grid ($n_x = 6$ grid cells) and is refined up to $n_x = 24$, using the $\kappa$ estimated at each previous step as the prior. For the double-peak system we start with $n_x = 10$ and refine up to $n_x = 40$. For both systems the final reconstruction contains less than one particle per grid cell.

For the PBL reconstruction, we use a smoothing scale of the form

$$h_n = \frac{c}{(\rho_0)^2},$$  \hspace{1cm} (39)

where $\rho_0$ is the local number density of points, $c$ is a constant, and $\xi$ is a tunable parameter to maximize the signal-to-noise ratio. For our simulation $\xi = 1$ is an optimal choice, and for the observational case, we have used $\xi = 0.5$, which is a common choice for equalizing the signal-to-noise ratio for every smoothing length. We select $c$ such that the integrated signal-to-noise ratio is greater than unity. The PBL reconstructions are gridded to the same resolution as the grid-based reconstruction to aid visualization.

For both reconstructions, we begin our iterations with a best-fit SIS. We do not, however, use this in the regularization for the grid-based reconstruction.
TABLE 1
Comparison between PBL and Grid-based Method

| Method       | Galaxy | Local Density | Grid Cells | $\chi^2$/DOF | $\eta$ | $c$ |
|--------------|--------|---------------|------------|--------------|-------|-----|
| **Single Peak** |        |               |            |              |       |     |
| PBL .......... | 0.0200 | 0.0136        | 0.0147     | 1.03         | 0.5   |     |
| PBL .......... | 0.0181 | 0.0128        | 0.0119     | 0.94         | 0.7   |     |
| PBL .......... | 0.0219 | 0.0139        | 0.0131     | 0.95         |       |     |
| PBL .......... | 0.0235 | 0.0140        | 0.0133     | 0.94         | 1.3   |     |
| PBL .......... | 0.0227 | 0.0120        | 0.0121     | 0.98         | 1.5   |     |
| Grid .......... | 0.0311 | 0.0283        | 0.0237     | 0.6          | 10    |     |
| Grid .......... | 0.0309 | 0.0280        | 0.0223     | 0.79         | 30    |     |
| Grid .......... | 0.0311 | 0.0280        | 0.0224     | 0.94         | 60    |     |
| **Double Peak** |        |               |            |              |       |     |
| PBL .......... | 0.0250 | 0.0174        | 0.0167     | 0.82         | 1.1   |     |
| PBL .......... | 0.0231 | 0.0168        | 0.0160     | 0.80         | 1.38  |     |
| PBL .......... | 0.0277 | 0.0193        | 0.0180     | 0.82         | 1.7   |     |
| PBL .......... | 0.0320 | 0.0219        | 0.0208     | 0.87         | 2.0   |     |
| Grid .......... | 0.0570 | 0.0711        | 0.0630     | 0.92         | 20    |     |
| Grid .......... | 0.0367 | 0.0497        | 0.0399     | 0.7          | 40    |     |
| Grid .......... | 0.0359 | 0.0482        | 0.0454     | 0.83         | 60    |     |

Notes.—The second, third, and fourth columns represent the deviation of the reconstructed $\kappa$ from the true $\kappa$. The second and third columns are weighted uniformly by galaxy, $\sum_{i=1}^{N} (\kappa_i - \kappa_{\text{model}})^2/N_{\text{gal}}$, and weighted by local density within each grid cell, $\sum_{i=1}^{N} (\kappa_i - \kappa_{\text{model}})^2/N_{\text{gal}}$. In the fourth column, equal weighting is given to every grid cell.

5.1.3. Results

Before discussing the results, we note a potential complication. Gravitational lensing mass measurements suffer from the mass sheet degeneracy when the sources are not distributed in redshift. This implies that $\kappa$ can be determined up to a degeneracy $\lambda\kappa + (1 - \lambda)$. This transforms to a degeneracy in the potential of the form

$$\psi(\theta) \rightarrow \psi'(\theta) = \frac{1}{2} (1 - \lambda) \theta^2 + \lambda \psi(\theta).$$

For the simulated data we have computed the best value of $\lambda$ in each case and transformed our reconstruction with that value of $\lambda$ for both the grid-based method and PBL.

In Table 1 we compare the $\chi^2$ values for the best fits of both the grid-based reconstruction along with PBL for a variety of smoothing normalization parameters, $c$. The aim of this table is to quantify the deviation of the reconstructed $\kappa$ from the true $\kappa$. In each case, the ostensible $\chi^2$/DOF is of order unity. However, one needs to be careful with simply asserting that the lower $\chi^2$ produces the best result, since the regularization in grid-based reconstruction adds a penalty function, and the smoothing scale in PBL lowers the effective degrees of freedom.

So while both models produce small values of $\chi^2$, the real question is whether these good fits correspond to an accurate reconstruction of the underlying density field. In Table 1 we do several comparisons which relate the reconstructed $\kappa$ at each galaxy (or grid point) with the true $\kappa$ modeled by the simulation for both the single peak and the double peak. The comparisons are done with a range of values for both $\eta$ (the regularization weight in the grid-based method) and $c$ (the proportionality constant in PBL).

The first column in Table 1 describes the method used, i.e., either PBL or the grid-based method. The second column describes the difference between reconstructed $\kappa$ and the true $\kappa$ for every galaxy position. In order to extract this information from the gridded reconstruction we have used the nearest grid point method, which simply means the $\kappa$ at each galaxy position is assigned the value at the corresponding grid cell. The third column describes the deviation of the reconstructed $\kappa$ from the true $\kappa$ at every grid cell weighted by the number of image galaxies in that grid cell. The fourth column describes the difference between the reconstructed and true $\kappa$ weighted uniformly over the grids. In each of the three comparisons, PBL reproduces the original reconstruction with the highest fidelity. The fifth column gives the $\chi^2$/DOF, the sixth gives the regularization parameter $\eta$ for the grid-based method, and the seventh column gives the smoothing normalization parameter for PBL.

In Figure 2 we show the radial reconstruction of the softened isothermal sphere using the two different techniques. The bulk of the penalty associated with the grid-based reconstruction relative to PBL occurs near the core. By construction, PBL is designed to perform well in this regime.

5.2. Simulation: A Double-Peaked Cluster

5.2.1. Model

While PBL has been shown to perform well modeling a single softened isothermal sphere in §5.1, the other major goal of this method is to reconstruct small-scale substructure in a system. To that end, we model a doubly peaked system with 814 lensed background galaxies. As before, they are placed on a unity grid and are modeled as two softened isothermal spheres, with $x_1 = 0.65$, $y_1 = 0.35$, $x_2 = 0.35$, $y_2 = 0.65$, $\theta_{E1} = 0.2$, $\theta_{E2} = 0.2$, and $\theta_{E2} = 0.1$. The simulated noise and reconstruction technique for the double-peaked system are identical to the single-peak system. This is a system of two subclusters at a redshift of $z_{\text{rms}} = 0.4$ having a velocity dispersion of $\sigma_v = 850$ km s$^{-1}$ separated by 226 kpc. The field of view is 105$''$.

5.2.2. Results

As with a single sphere, both PBL and grid-based reconstructions produce $\chi^2$/DOF $\approx 1$, as illustrated in Table 1. However, as with the single-sphere reconstruction, PBL produces smaller errors with regards to the underlying model than does the grid-based reconstruction.

In Figure 3 we show a gray-scale plot of the residuals between the underlying model and each of the reconstructions. Unsurprisingly, both models have the greatest difficulty reproducing highly peaked cores, although PBL is more responsive to high local gradients in $\kappa$. We describe the general quality of the fit in Table 1.

5.3. Observation: The Bullet Cluster

5.3.1. Observations

Finally, we perform a mass reconstruction of the Bullet Cluster (1E 0657–56). This galaxy cluster is a rare supersonic merger in the plane of the sky. Its distinctive structure and orientation make it an ideal cluster for observing dark matter using gravitational lensing. It consists of two subclusters separated by 0.72 Mpc, which have just undergone a merger and are moving away from each other. The western subcluster is less massive, and the eastern main cluster is more massive. The line-of-sight velocity difference suggests that their cores passed each other 100 Myr ago. The collisionless dark matter in each of the subclusters have crossed each other, but the fluidlike intracluster plasma is in the process of electromagnetic and thermal interaction, producing high X-ray luminosity far removed from lensing mass peaks (Clowe et al. 2006; Bradac et al. 2006).
For the Bullet Cluster, we perform a PBL reconstruction only, since it has been well studied with grid-based methods (using Schneider 1995; Kaiser 1995) and the $\kappa$-contours are publicly available. We use publicly available weak-lensing data from the Bullet Cluster Project Page. The catalog was constructed using data from three different instruments: the ESO/MPG Wide Field Imager, IMACS on Magellan, and two pointings of ACS on HST. The shapes of the galaxies were measured independently on each of the image sets, averaging for the common galaxies. The weighting for each galaxy is based on its significance of detection in every image set and is normalized appropriately (Clowe et al. 2006).

The catalogs were combined using weighted average reduced shear measurements, and the weights of individual galaxies were increased when they occurred in several catalogs. This weighting is listed in the shear catalog. We include this weighting in our reconstructions as well and choose only those images with a weighting greater than 1. As we have already illustrated in the simulations, PBL is most effective when the information density is variable, i.e., close to the core of the clusters. In the case of the Bullet Cluster we

---

2 See http://flamingos.astro.ufl.edu/1e0657/public.html.
zoom into a region bounded by $104.53^\circ - 104.69^\circ$ in right ascension and $55.92^\circ - 55.97^\circ$ in declination. Following this cut, our sample includes 1259 weak-lensing background galaxies. In order to do the mass reconstruction we use the average redshift of this sample, $z = 0.91$.

5.3.2. Reconstruction

The Bullet Cluster was made famous by the direct detection of dark matter by Clowe et al. (2006). Indeed, since one of the major findings of this group is that the dark matter appears offset from X-ray emissions, we do not include any prior model when reconstructing the system, but are able to achieve fast convergence with two clearly visible peaks. This reconstruction guides us in choosing an initial condition for subsequent $\chi^2$ minimization.

We have calculated the integrated mass within 150 kpc of each peak. The main peak has a mass of $1.57 \times 10^{14}$ $M_\odot$, and the subcluster has a mass of $0.9 \times 10^{14}$ $M_\odot$. Clowe et al. (2004) report a value of $(1.02 \pm 0.16) \times 10^{14}$ $M_\odot$ for the main peak and $(0.66 \pm 0.19) \times 10^{14}$ $M_\odot$ for the subcluster within 150 kpc of each peak. In each case, our estimate exceeds that of Clowe et al. by approximately 3.4 $\sigma$. However, a more recent S+W reconstruction by the same group (Branch et al. 2006) yields masses of $(2.8 \pm 0.2) \times 10^{14}$ $M_\odot$ around the main peak and $(2.3 \pm 0.2) \times 10^{14}$ $M_\odot$ around the subcluster within 250 kpc of each peak. Inclusion of strong-lensing information makes reconstruction of the cores more accurate and also leads to higher estimates of the mass. Even correcting for the greater area, this suggests Clowe’s initial mass estimate may have been low.

Our mass estimates using PBL are higher than the weak-lensing reconstruction of Clowe et al. (2004) and, thus, more in line with the S+W results. This is a result of a difference in method. For example, we start from an initial condition and iterate to the correct solution, whereas Clowe et al. (2004) have fitted a radially averaged shear profile to the NFW or King profile. As already seen in the simulations, using an initial condition recovers values of $\kappa$ close to the core with greater accuracy. This implies that while most weak-lensing $\kappa$-maps report $\kappa$-contours less than 1, using an initial condition and PBL we are able to get $\kappa > 1$ greater than 1. This implies that the mass we measure will also be greater than the typical weak-lensing mass measurement. Also to measure the mass of the subcluster, Clowe et al. (2004) have removed the mass of the main cluster to avoid overestimation of the mass; we have not considered this effect in our reconstruction.

In Figure 4 we show our PBL reconstruction of the Bullet Cluster. Note that, despite using weak-lensing signals only, we are able to identify both density peaks, and using initial conditions, we are able to get $\kappa > 1$ for the main peak. We also do a comparison of the publicly available $\kappa$-contour with the $\kappa$-contours reconstructed using PBL. The location of the main peak coincides for both reconstructions. The subcluster contours for PBL are slightly removed from the publicly available $\kappa$-contours.

Error analysis for PBL will be discussed in detail in future papers. In particular, the noise covariance matrix, $\langle (\kappa - \kappa) (\kappa - \kappa)^T \rangle$, will give us important insights into the errors caused by the reconstruction method. A bootstrap method can also be used to determine error bars on mass measurements from observations. In the case of simulations, several Monte Carlo realizations of the noise can be used to study the errors.

6. DISCUSSION AND FUTURE PROSPECTS

6.1. Additional Signals

Thus far, we have developed the formalism for PBL and done worked examples demonstrating how it may be applied to weak-lensing reconstruction. It is designed to model structure hierarchically, in part because of the great success of strong+weak lensing analysis.

Several groups have already shown how multiple image positions may be added to the information yielded by lens ellipticities to produce very high quality mass maps of clusters. It was our desire to maximally exploit the different information scales of the strong- and weak-lensing signals which motivated the development of PBL in the first place.

However, there is yet more information besides image differences potentially available which may be utilized in a reconstruction. Consider that in addition to the two constraints generated by the positional difference between two images, we also can measure a flux ratio and two ellipticity differences. Thus, in principle, we have five measurable model parameters per strong-lensing pair rather than two, and in an idealized case, this improves the potential resolution of a system in the strong-lensing regime by $(5/2)^{1/2} \approx 1.6$. As a way of guiding the future development of PBL, we discuss possible future avenues of investigation below.

6.1.1. Flux

Apart from the centroid position, the Petrosian flux of an image is the most straightforward to measure. The magnification of the lens is given by the inverse of the determinant of the projection matrix,

$$\mu = \frac{1}{(1-\kappa)^2 - |\gamma|^2}.$$ 

(41)
Unlike the displacement vectors \( \alpha \), which are simple linear operators of the potential field (the gradient), or the weak-lensing shear field, which is nearly so (since in the limit of \( \kappa \ll 1 \), the image ellipticity is an unbiased estimator of the shear field), the flux is a highly nonlinear function of the shear and convergence fields. This accounts, in part, for the reason that it has not been used previously in cluster reconstructions. Here we note that Saha et al. (2007) show that the image positions themselves constrain the fluxes for a source with three noncollinear components. This is a special case, for cluster lensing three-component sources for strong-lensing may not always be available. Also, Natarajan et al. (2007a) use magnification information in their parametric mass modeling of clusters.

The other major consideration is that magnification is not a smoothly varying function of the potential fields. It is well known that on the critical curves, magnification goes to infinity (see, e.g., Schneider et al. 1992 for an extensive discussion), but this is a set of measure zero, so in and of itself it produces no problem. The issue is that the parity of the image reverses as an image that crosses the critical curve.

Negative magnification means nothing more than reversal of image parity and, thus, cannot generally be easily detected. Thus, we are much more interested in computing terms which scale like \( \kappa^2 \). Indeed, since we cannot measure the magnification directly, but only the flux, we propose that the combination

\[
\frac{\mu_A^2 - \mu_B^2}{\mu_A^2 + \mu_B^2} = \frac{f_A^2 - f_B^2}{f_A^2 + f_B^2}
\]

is directly measurable and has no poles.

Even so, a lensing model predicts a parity for a particular image, and as with ellipticity minimization, there is a discontinuity in the derivatives. In Figure 5 we show the magnification (including sign) as a function of convergence and shear.

### 6.1.2. Ellipticity Differences

Likewise, while most measurements of the shear are based on an assumption that any given image is randomly oriented, two images of the same source are not. The difference in their measured ellipticity can be wholly modeled by the relative lensing fields at their respective locations. If both images were in the weak regime, we would be able to use the simple estimator

\[
\varepsilon_A - \varepsilon_B = \gamma_A - \gamma_B, \quad (43)
\]

where all terms in the equation are complex, and thus provide two constraints with high signal-to-noise ratio per image pair.

In general, however, a more likely configuration is that one image may be in the strong regime and one in the weak. If we can determine from the configuration of lenses which is which, we might imagine a better estimator as

\[
\varepsilon_A - \varepsilon_B = \frac{1}{g_A} - g_B, \quad (44)
\]

with the only associated noise corresponding to photon noise rather than random variance in the intrinsic ellipticity of the images.

### 6.1.3. Flexion

Thus far, the analysis of clusters in the weak or semiweak regime has primarily relied on shear. However, recently, Okura et al. (2008) and Leonard et al. (2007) have worked on reconstructing A1689 using flexion. In particular, the Okura group used a Fourier inversion suggested by Schneider & Er (2008). However, the advantage of our proposed PBL is that flexion (and, in principle, any higher order derivative of the potential) may be explicitly included as an additional constraint in the cluster reconstruction. Unlike Fourier techniques, which rely on binning
of the data, the PBL method will allow us to exploit the natural small-scale signal probed by flexion.

6.2. Summary

We have developed PBL, a new particle-based technique of mass reconstruction of clusters. The distinguishing feature of PBL is its ability to adjust its smoothing scale depending on the local signal-to-noise ratio or the type of constraint and thus not require any regularization. PBL has the scope of calculating derivatives up to any order. Hence, lensing constraints that are a function of the derivatives of the potential can be easily included in the reconstruction. In this paper we have successfully applied PBL to do weak-lensing—only mass reconstruction for a single-peak and a double-peak system. We have made the codes for PBL publicly available for application to weak-lensing measurements through our Web site (see § 4). The codes have been tested on the data sets and simulations described in the paper. A larger data sample will require modification of the current version of code.

As already explained, PBL is a method of discretizing data and not a minimization method. A $\chi^2$ minimization does not necessarily ensure reaching a global minimum. In many cases the global minimum is guarded by steep walls surrounded by shallow valleys. Without any prior knowledge of the mass distribution it is very easy to get trapped in a shallow valley and not reach the global minimum. We have started with an initial condition and interpolated the ellipticity function to aid us in this regard.

In future work we will be including additional constraints, like the flux ratios, ellipticity differences, and flexion, along with measured ellipticities and strong-lensing positions. We will also be exploring different minimization schemes to facilitate convergence to a global minimum.

The authors would like to thank Mike Jarvis, Bhuvnesh Jain, and Kevin Olson for useful discussions. We would also like to thank Marusa Bradacˇ and Douglas Clowe for providing us with the “Bullet Cluster” (1E 0657–56) data. D. M. G. gratefully acknowledges discussions with Richard Massey. S. D. would like to thank Ben Metcalf, Prasenjit Saha, and Sudeep Das for valuable comments. We would also like to thank the referee for thoughtful suggestions. This work was supported by NASA ATP NNG 05-GF61G.