Noise Based Detection and Segmentation of Nebulous Signal

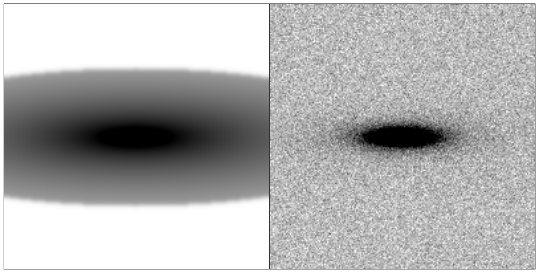
Mohammad Akhlaghi[1,2] & Takashi Ichikawa[1]

[1] Tohoku University Astronomical Insitute.
[2] GNU Astronomy Utilities maintainer.

ADASS XXV
26 October 2015

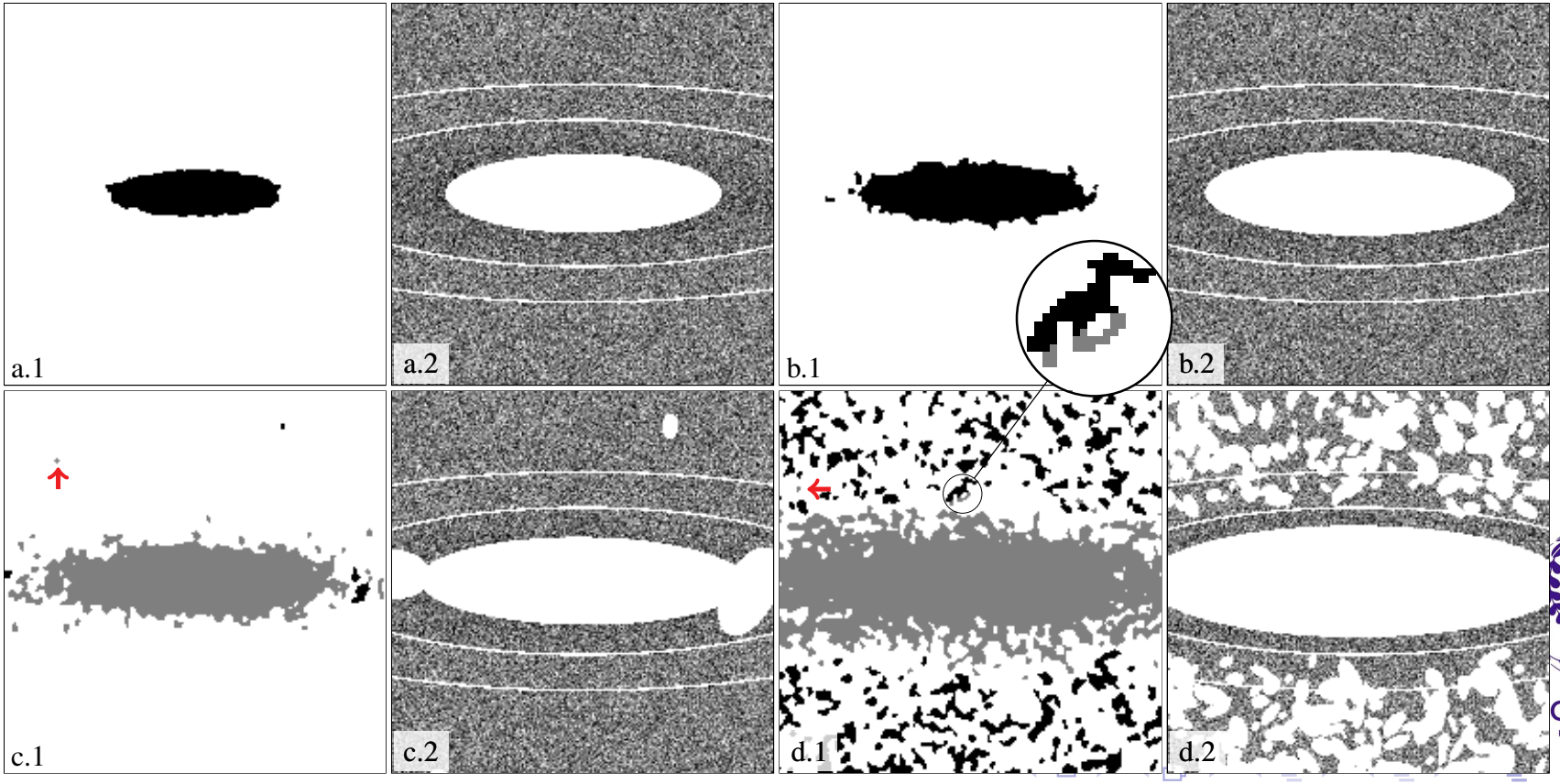
Growing and the Kron radius (r_k):

Kron (1980) radius: $r_k = \frac{\sum rI(r)}{\sum I(r)}$.

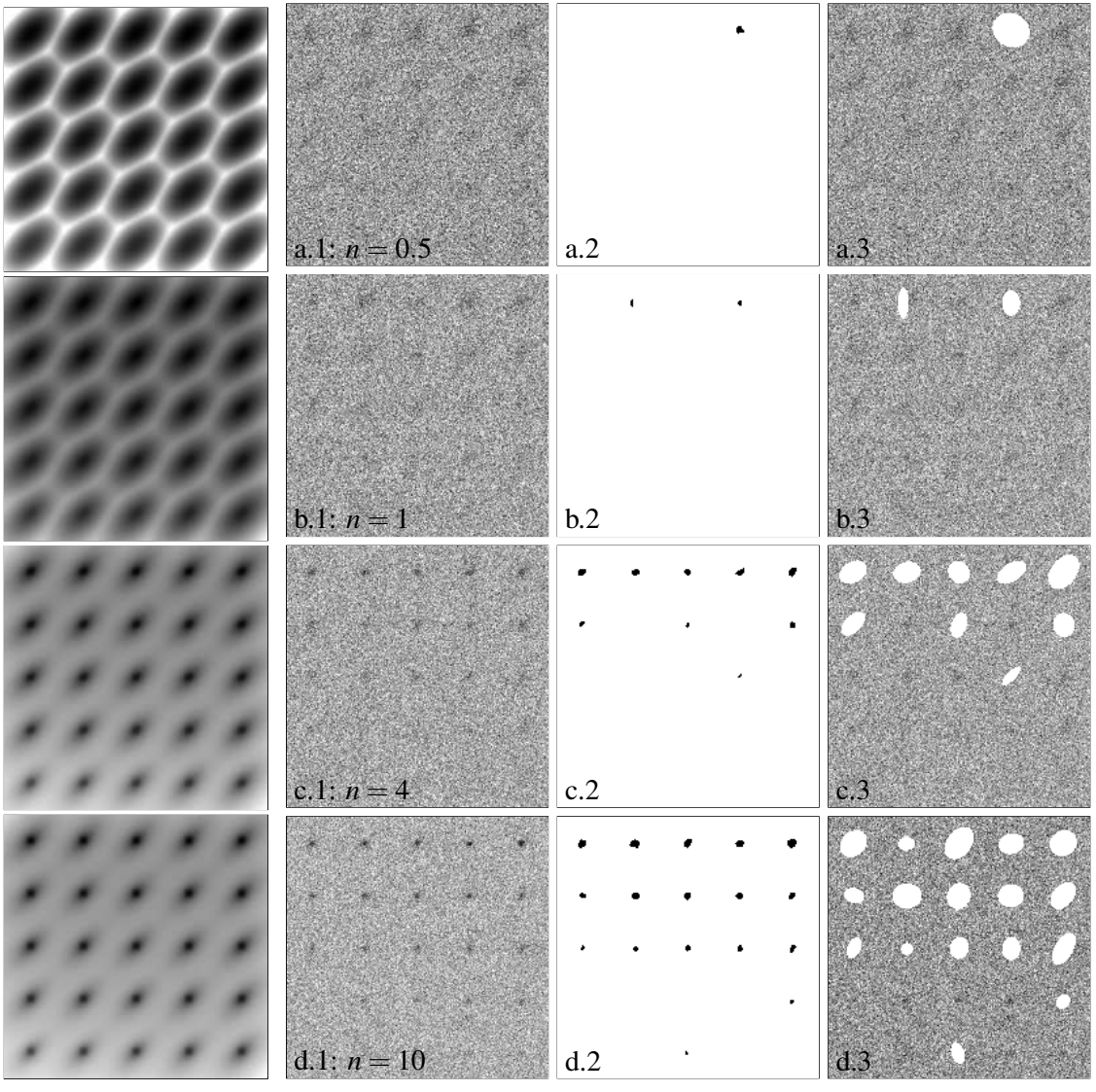


- Assumes an ellipse (very simplistic for real galaxies).
- Depends on the area above the threshold.
- Masked below: $3 \times r_k$ (hardly reaches 90% of flux).

`DETECT_THRESH` is 2, 1, 0.5 and 0.1 respectively.



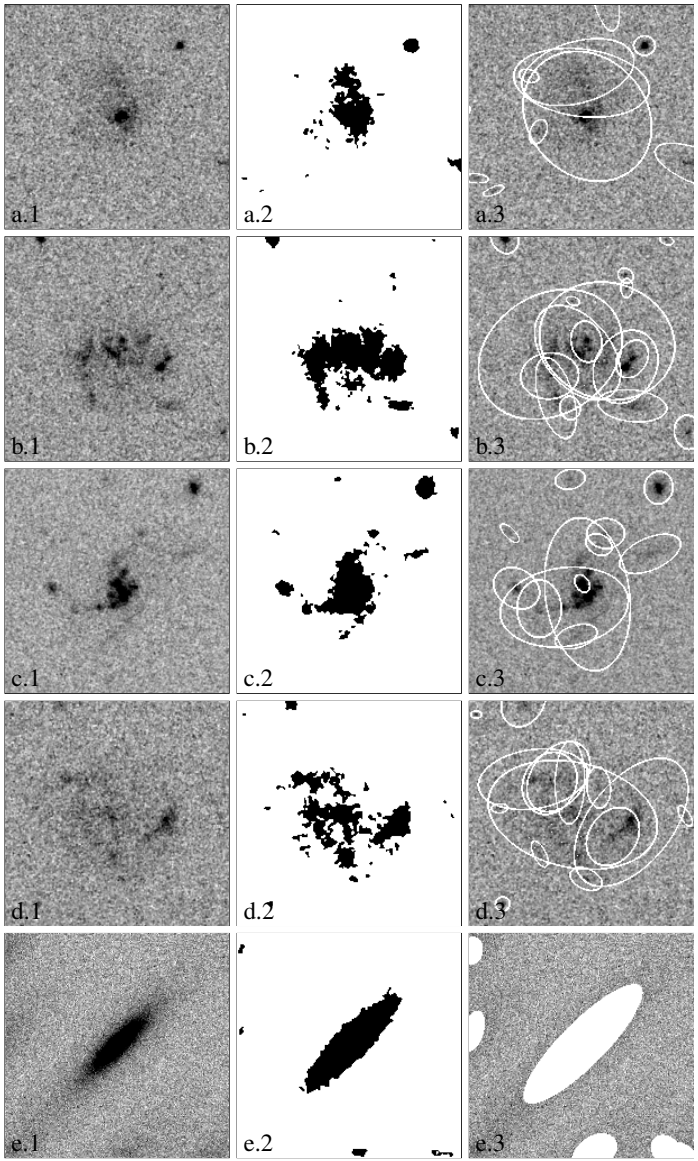
SExtractor: Sensitivity Test



GNU

TOHOKU UNIVERSITY

SExtractor: Real

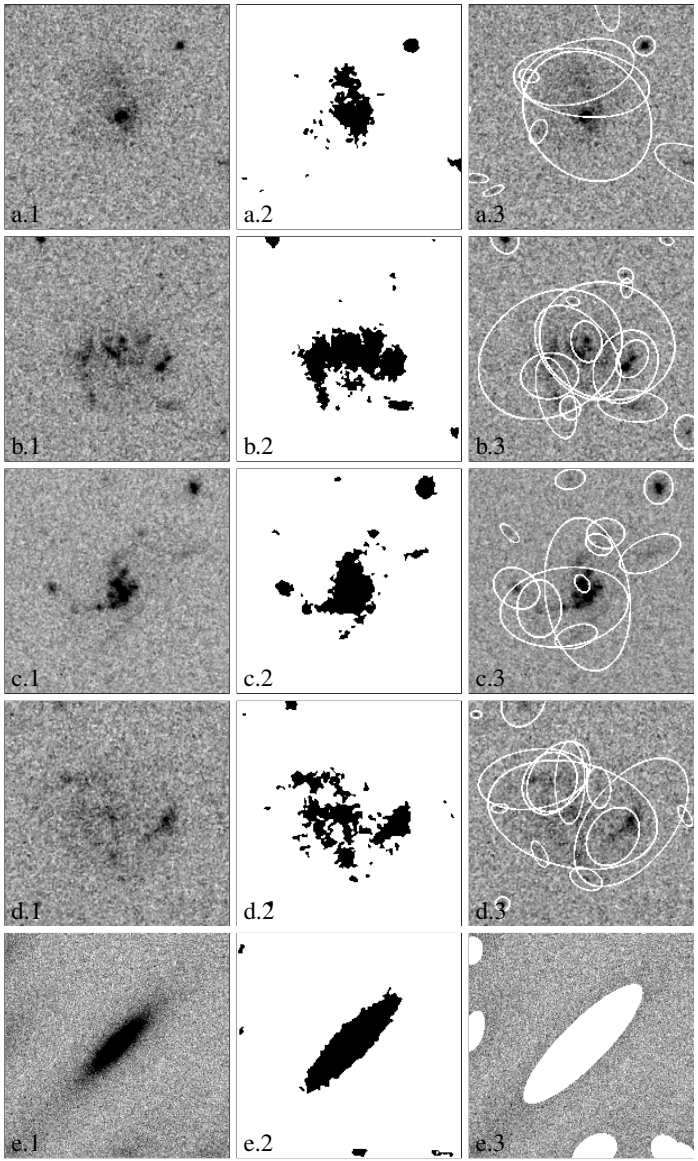


Real objects (images):

- Are not a clean ellipse.
  - Can be clumpy.
  - Can be diffuse.
  - Can have spiral arms.
  - Can be on the edge of the image.
- Do not necessarily have a uniform radial profile.

The existing method fails since such objects do not satisfy its *a priori* assumptions.

SExtractor: Real



Real objects (images):

- Are not a clean ellipse.
  - Can be clumpy.
  - Can be diffuse.
  - Can have spiral arms.
  - Can be on the edge of the image.
- Do not necessarily have a uniform radial profile.

The existing method fails since such objects do not satisfy its *a priori* assumptions.

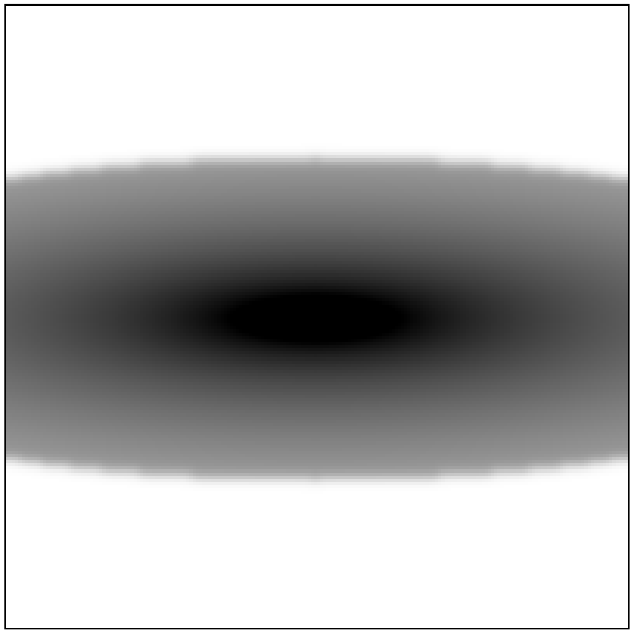
higher redshifts → more clumpy
(e.g., Murata et al. 2014)

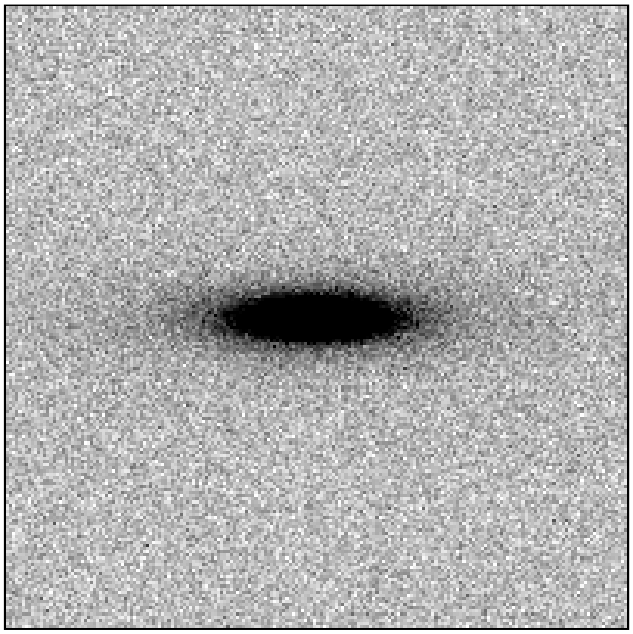
NoiseChisel – Detection – Basics

Aims:

- Threshold must be independent of the Sky.
- Impose negligible assumptions on signal.
- Accurately remove false detections.
- Use the actual data, not *a priori* models.

Model profile for demonstration:



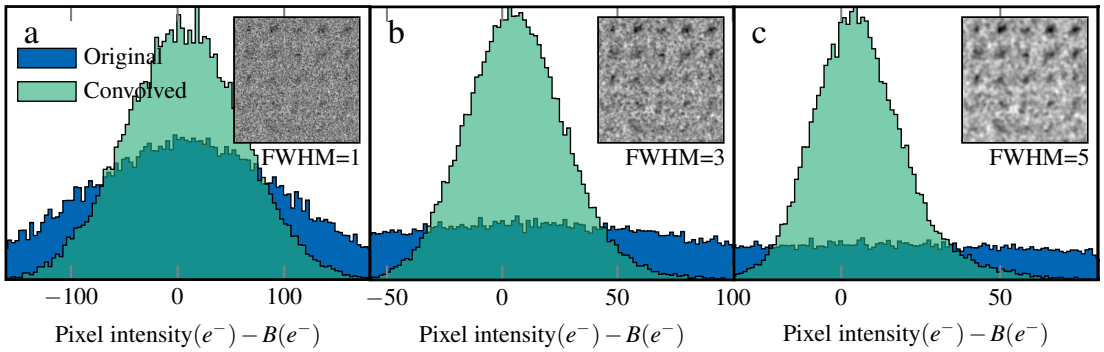


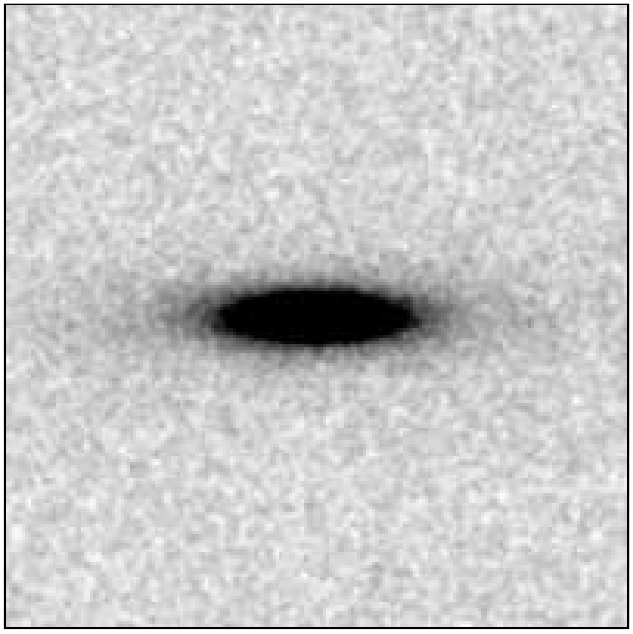
GNU

TOHOKU UNIVERSITY

NoiseChisel – Detection – Convolution

- Convolution decreases dynamic range.
- **So:** Gaussian kernel, FWHM= 2pixels.





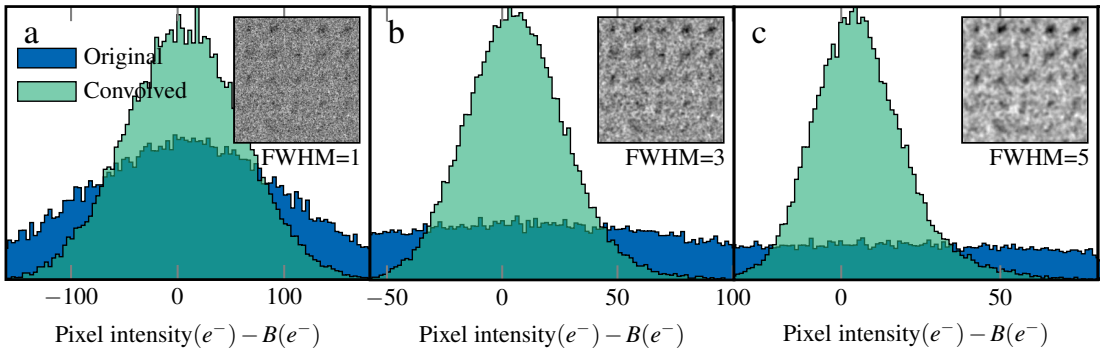
Since we are concerned with the sampling (noise) the PSF is no longer necessary and the same parameters work accuratly on space-based and ground-based images.

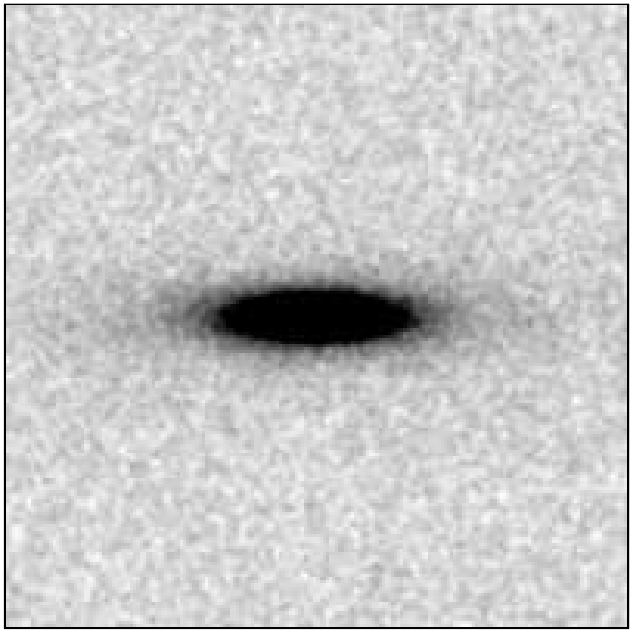
GNU

TOHOKU UNIVERSITY

NoiseChisel – Detection – Convolution

- Convolution decreases dynamic range.
- **So:** Gaussian kernel, FWHM= 2pixels.





Since we are concerned with the sampling (noise) the PSF is no longer necessary and the same parameters work accuratly on space-based and ground-based images.

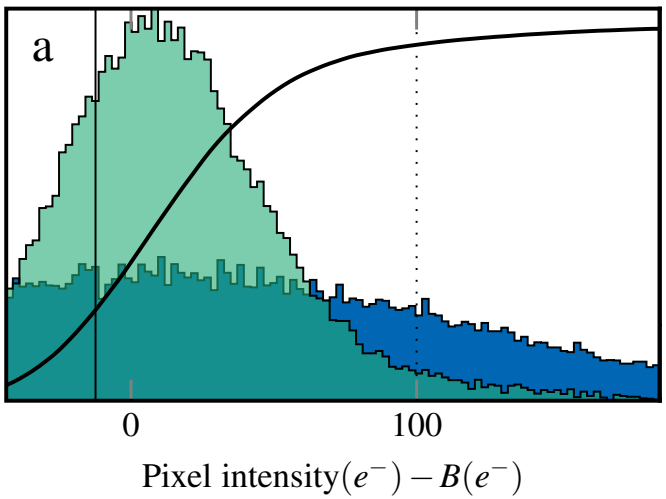
An assumption removed. Works on any image.

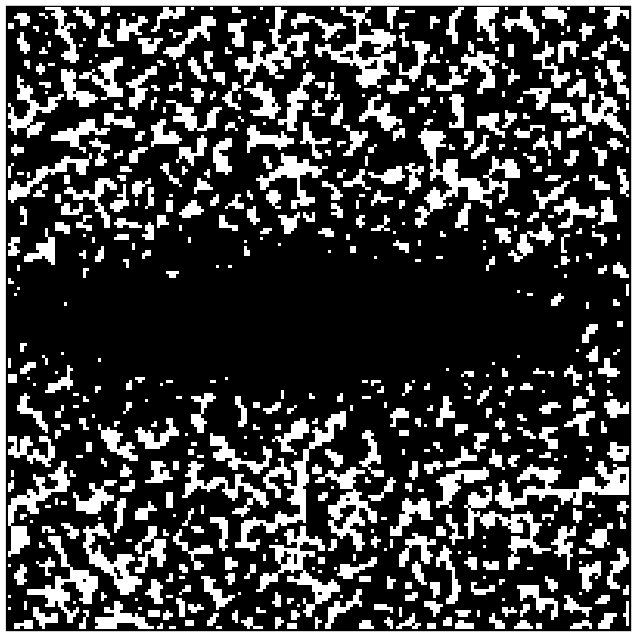
GNU

TOHOKU UNIVERSITY

NoiseChisel – Detection – Threshold

- Use the cumulative pixel distribution.
- The threshold is set to the 0.3 quantile of the image.





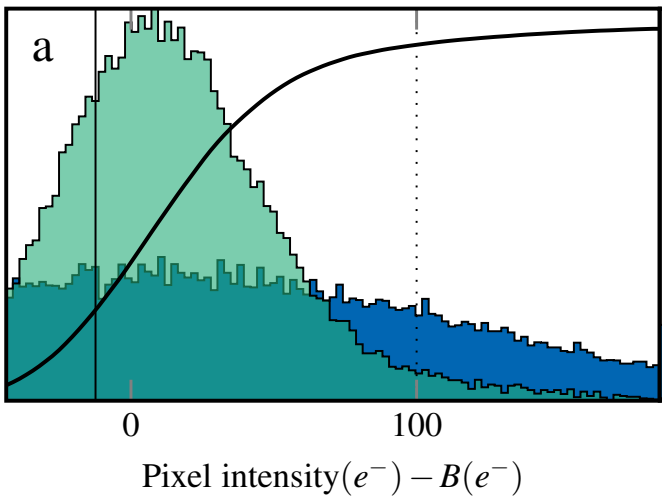
Since the threshold is now independent of Sky, we can accurately estimate the Sky once detection is complete.

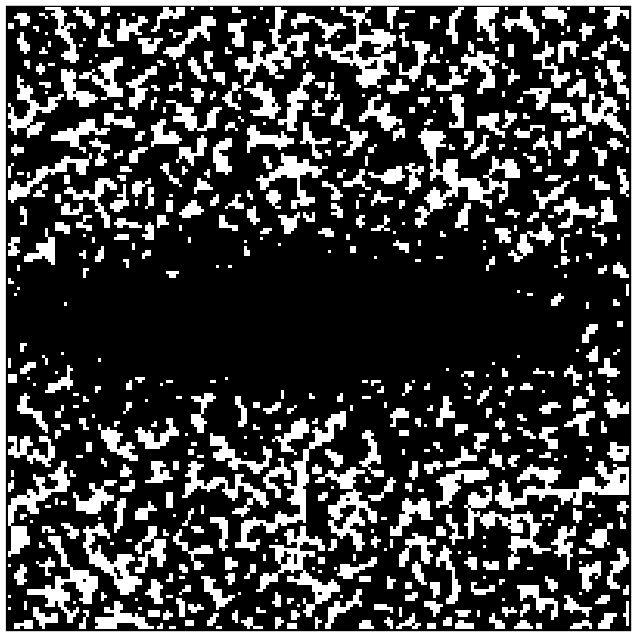
GNU

TOHOKU UNIVERSITY

NoiseChisel – Detection – Threshold

- Use the cumulative pixel distribution.
- The threshold is set to the 0.3 quantile of the image.





Since the threshold is now independent of Sky, we can accurately estimate the Sky once detection is complete.

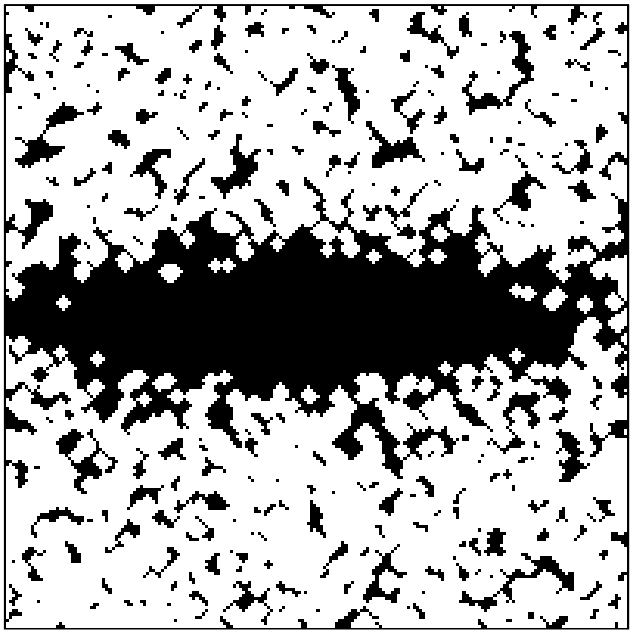
Threshold no longer defined by Sky.

NoiseChisel – Detection – Erode

Erosion: Foreground becomes background if touching.

- **Or:** we expand the holes.
- **Or:** we carve off the signal.

NoiseChisel name: a tool to carve off noise



GNU

TOHOKU UNIVERSITY

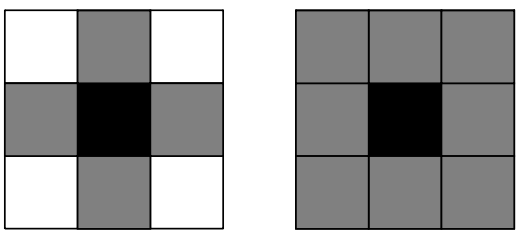
NoiseChisel – Detection – Open

Definitions:

- **Dilation:** Inverse of erosion.
- **Opening:** Erosion followed by dilation.

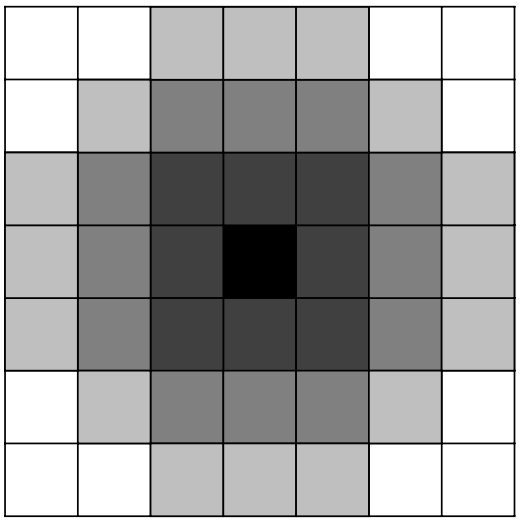
In practice:

- Separates all the steps below.
- We use eight connectivity here (and four connectivity in the previous step.)

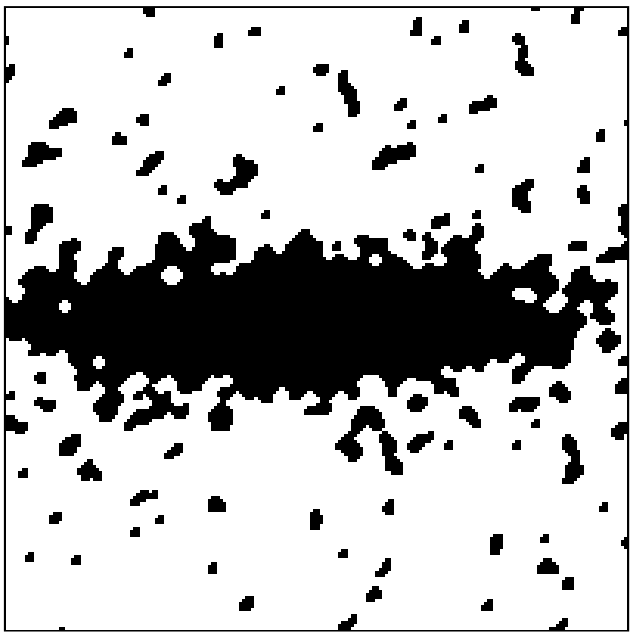


a

b



c

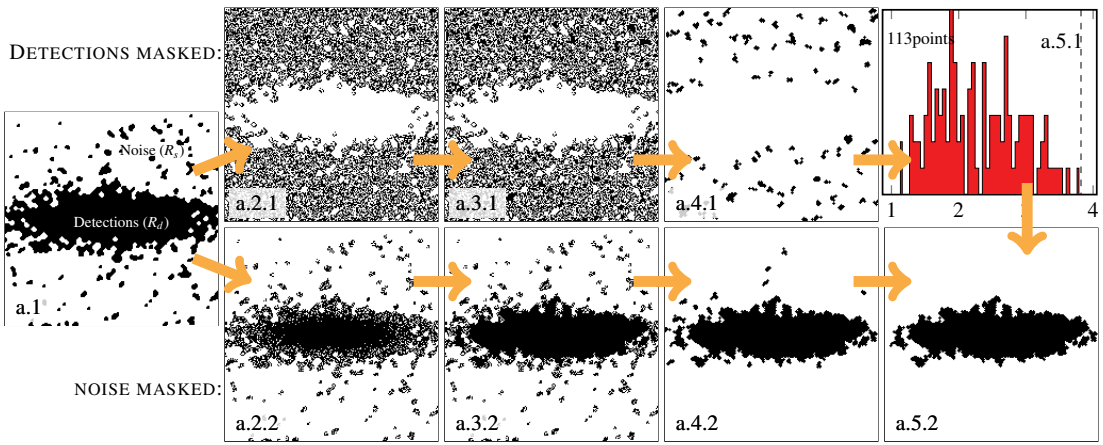


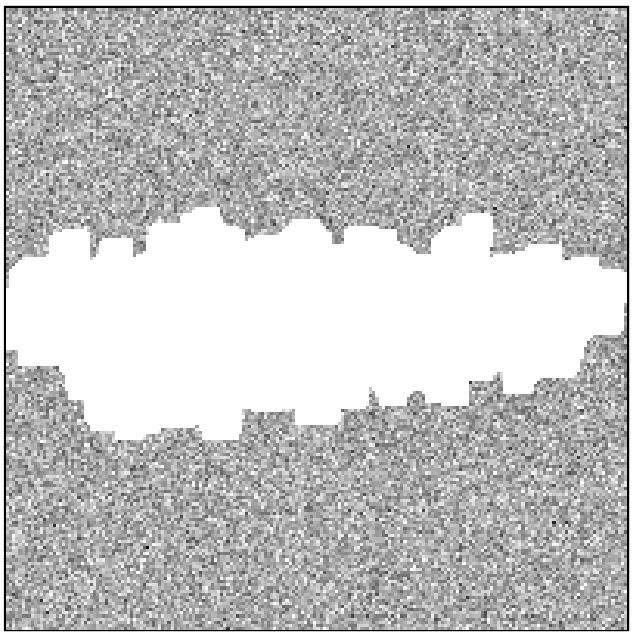
GNU

TOHOKU UNIVERSITY

NoiseChisel – Detection – Remove false detections

- Use the ambient noise as a reference.
- The S/N of definite false detections is used:





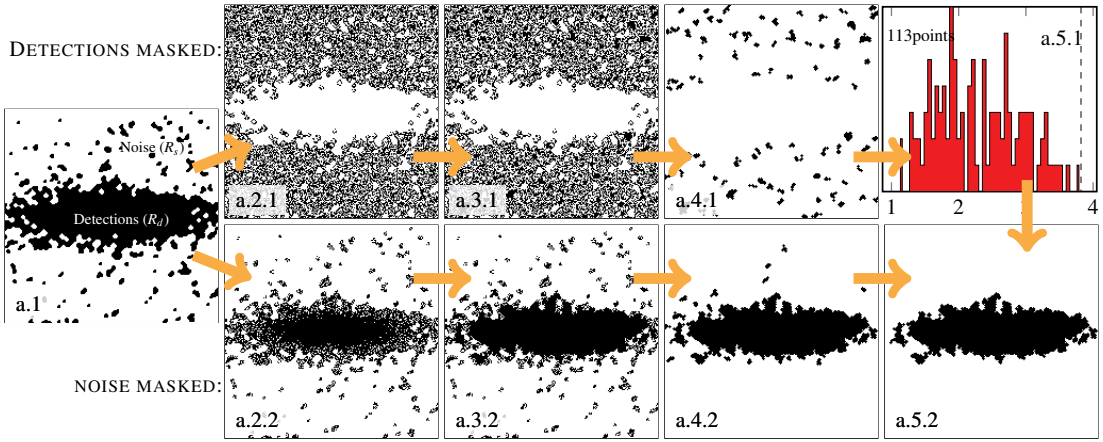
False detections are successfully removed with high accuracy.

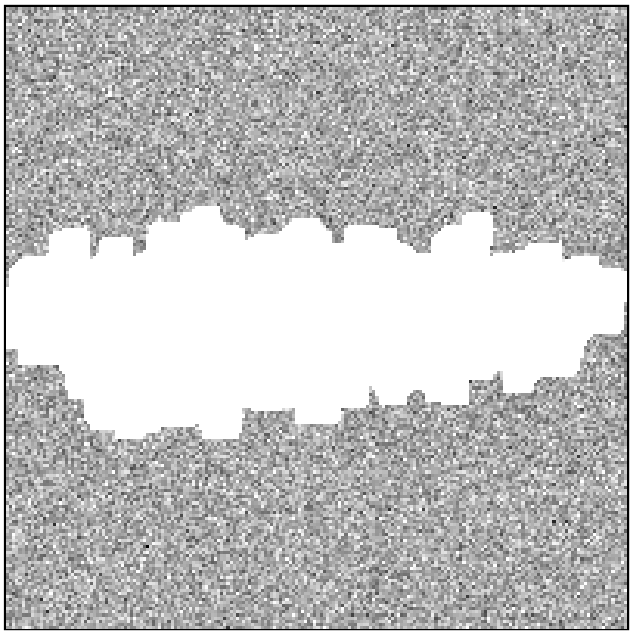
GNU

TOHOKU UNIVERSITY

NoiseChisel – Detection – Remove false detections

- Use the ambient noise as a reference.
- The S/N of definite false detections is used:





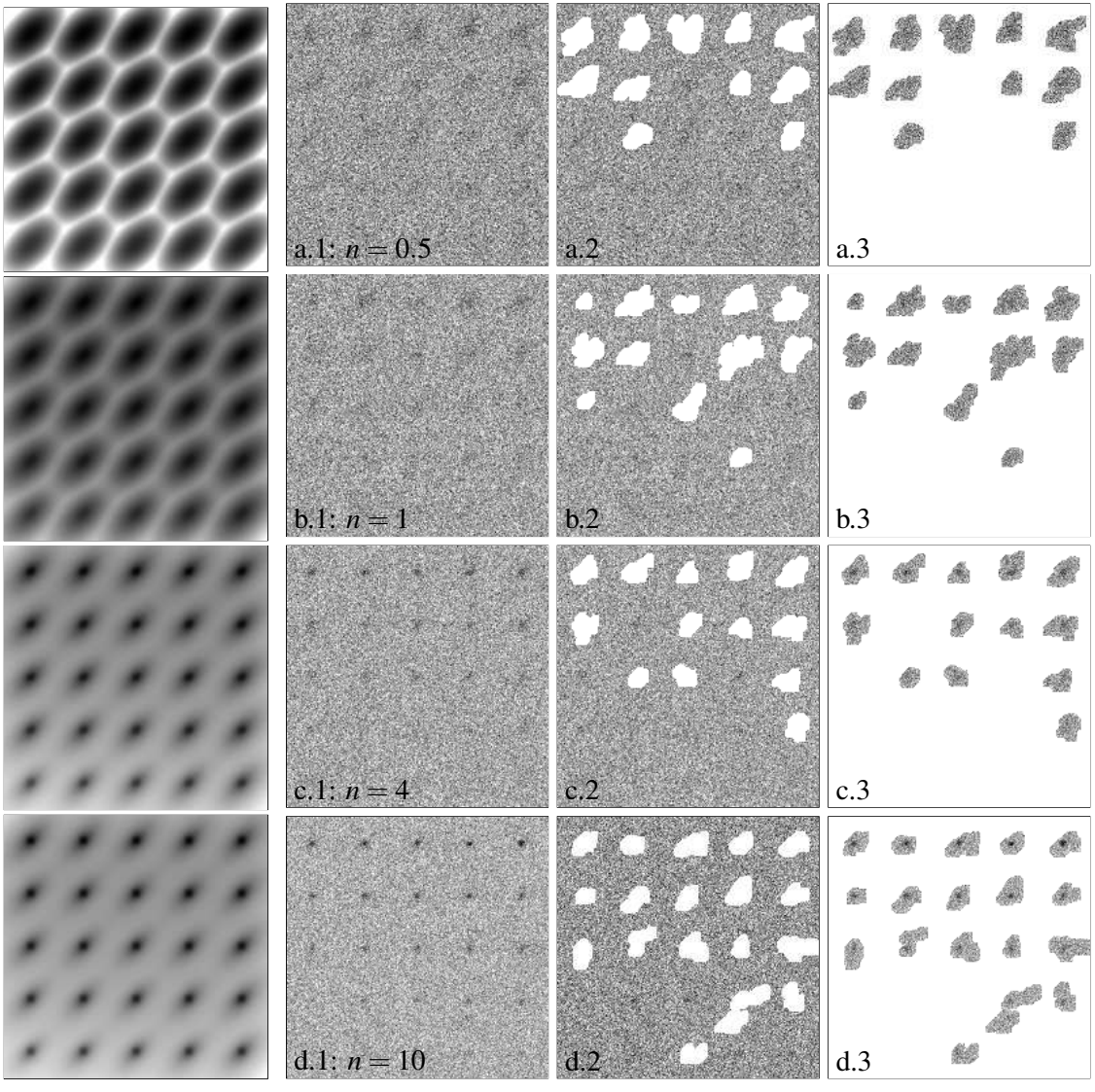
False detections are successfully removed with high accuracy.

✔ False detections are now identified for any image without hand-input values.

GNU

TOHOKU UNIVERSITY

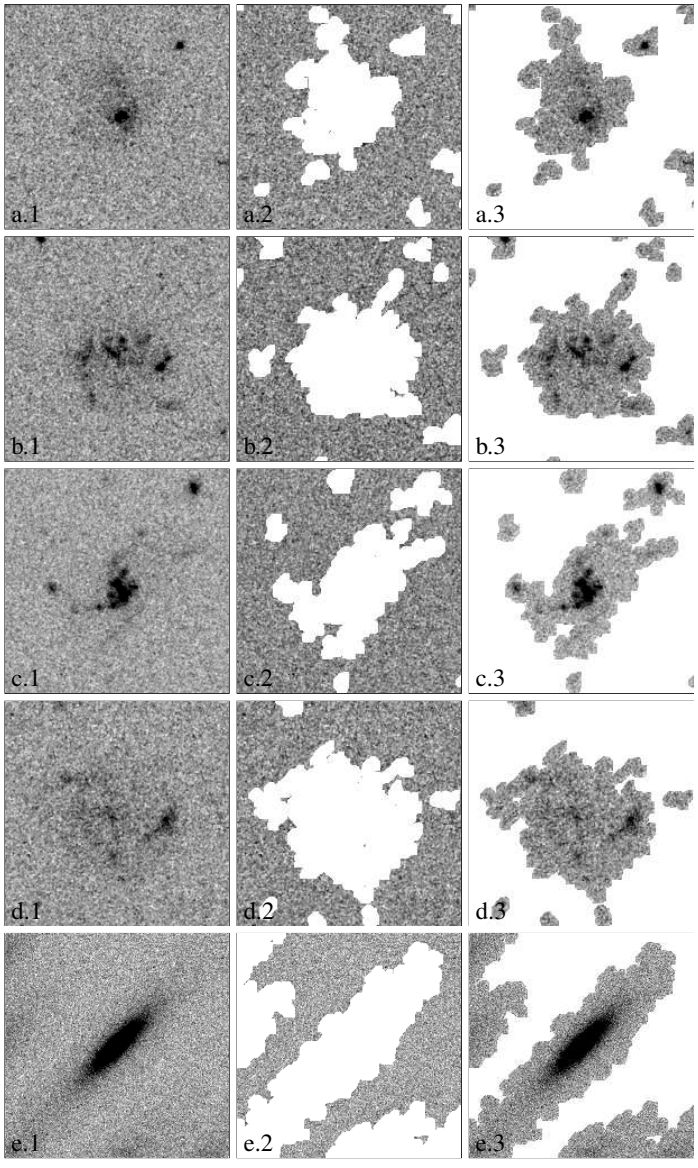
NoiseChisel – Detection – Sensitivity Test



GNU

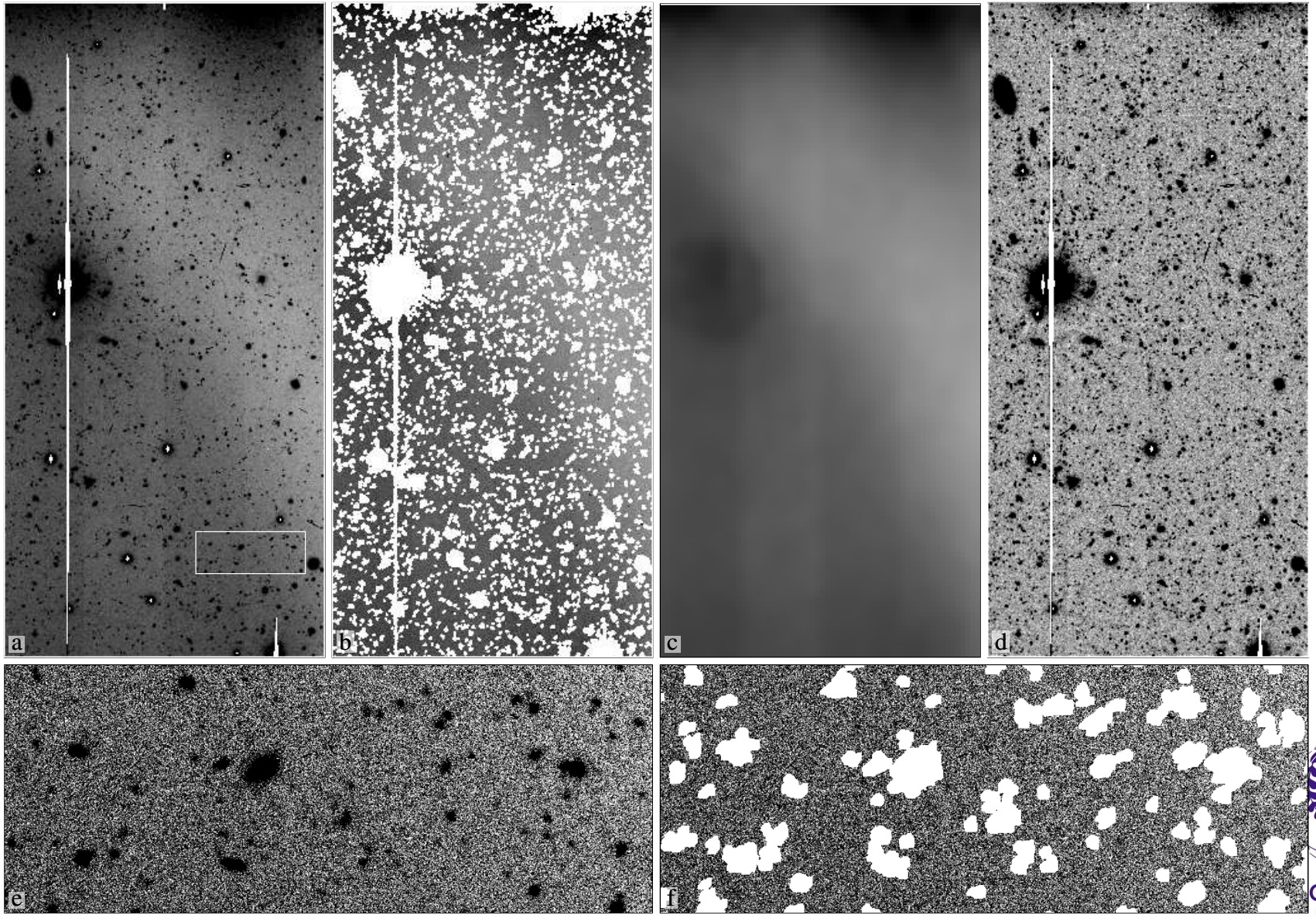
TOHOKU UNIVERSITY

NoiseChisel – Detection – Real images



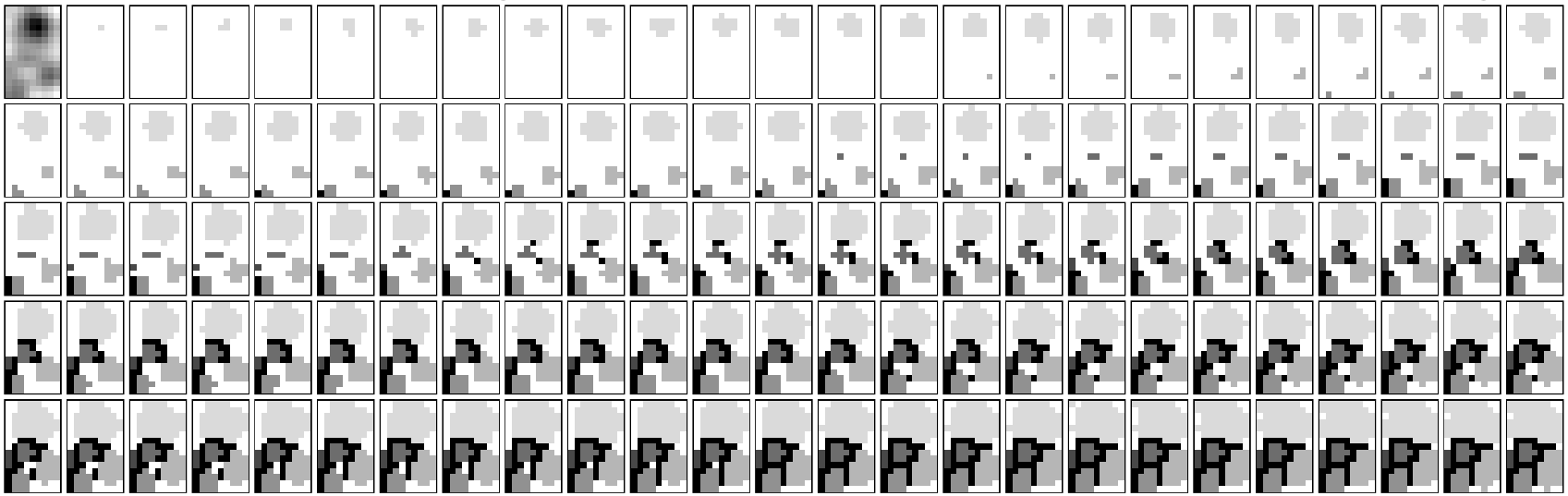
Noise-based detection: Works on any image with any target shape.

NoiseChisel – Detection – Large real images



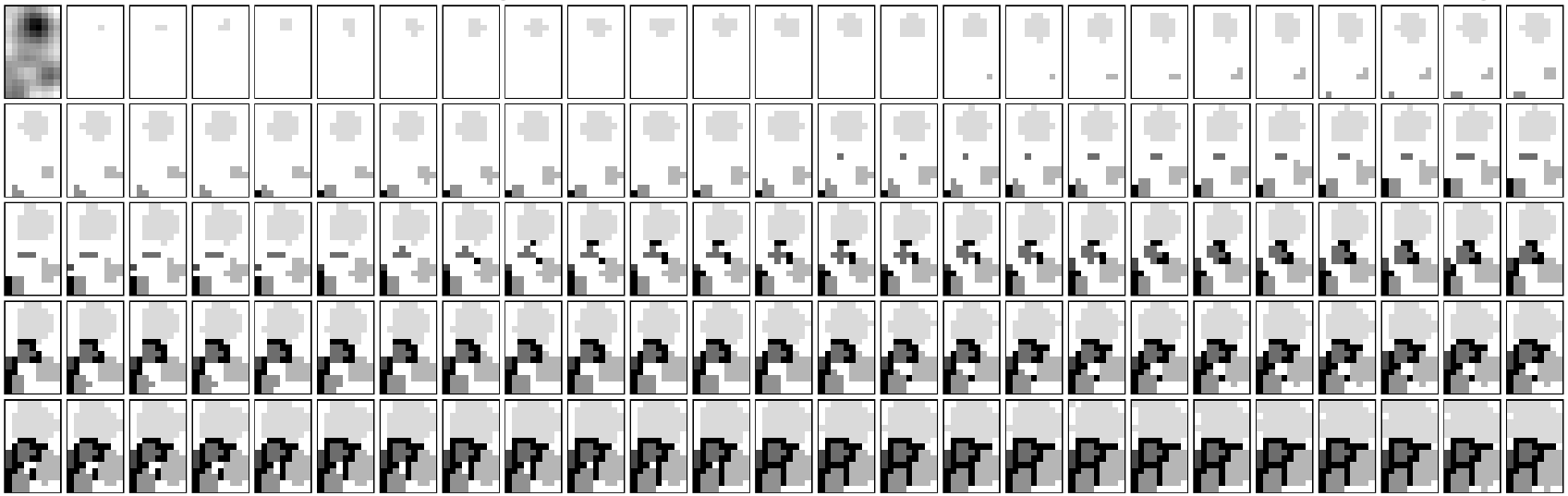
NoiseChisel – Segmentation – clumps

A clump is found using the maximum resolution of the convolved image:



NoiseChisel – Segmentation – clumps

A clump is found using the maximum resolution of the convolved image:

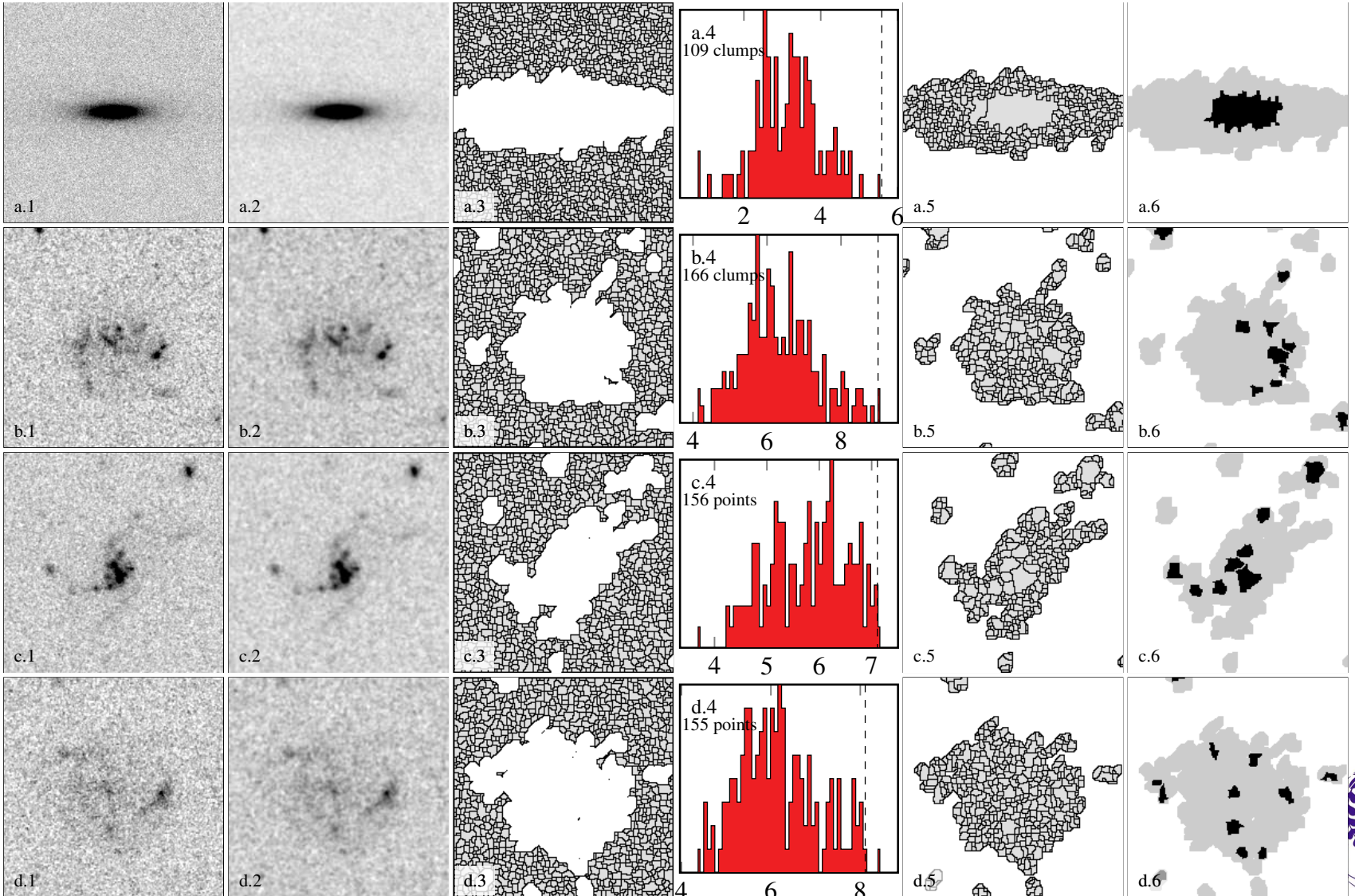


No more layers

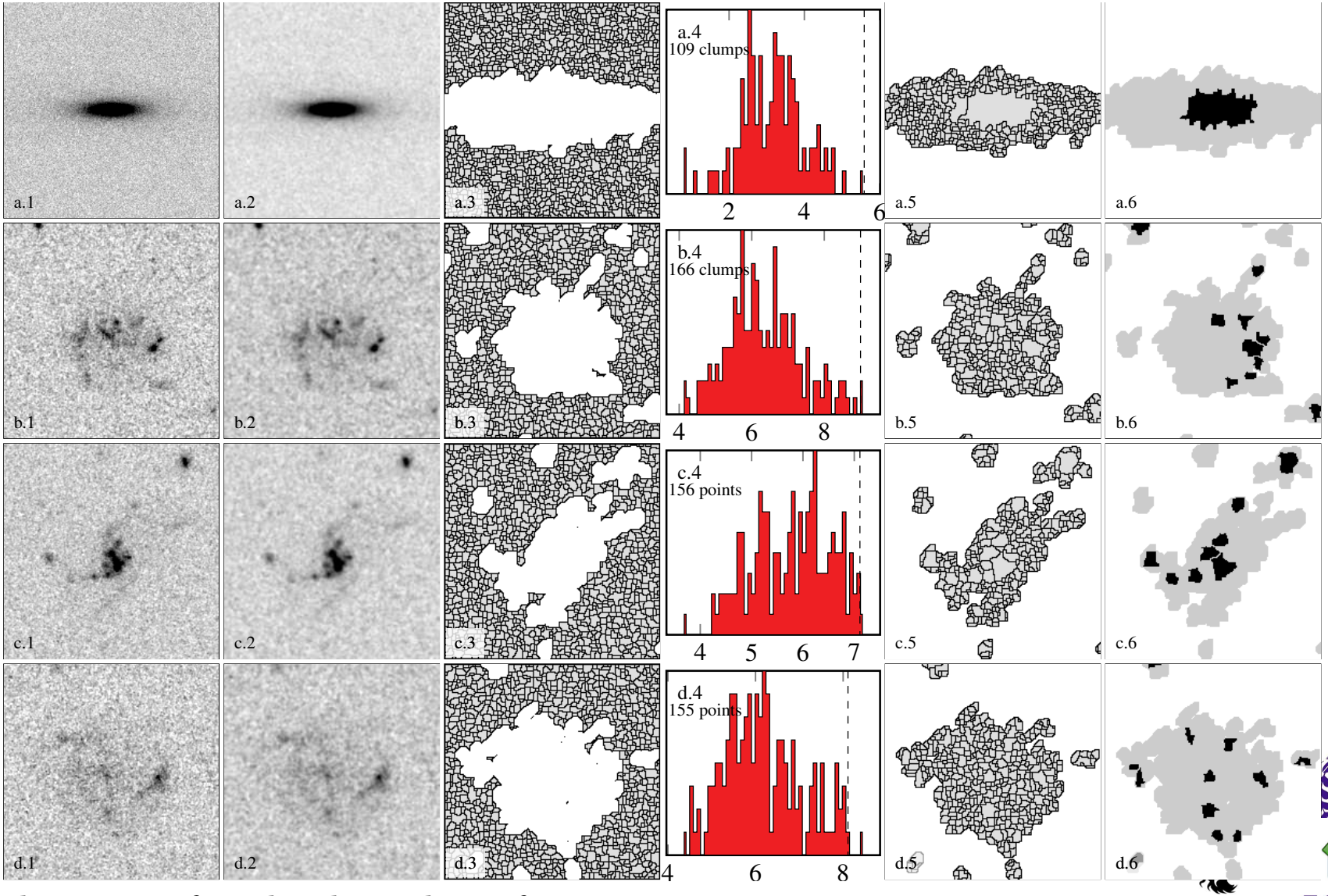
GNU

TOHOKU UNIVERSITY

NoiseChisel – Segmentation – True clumps

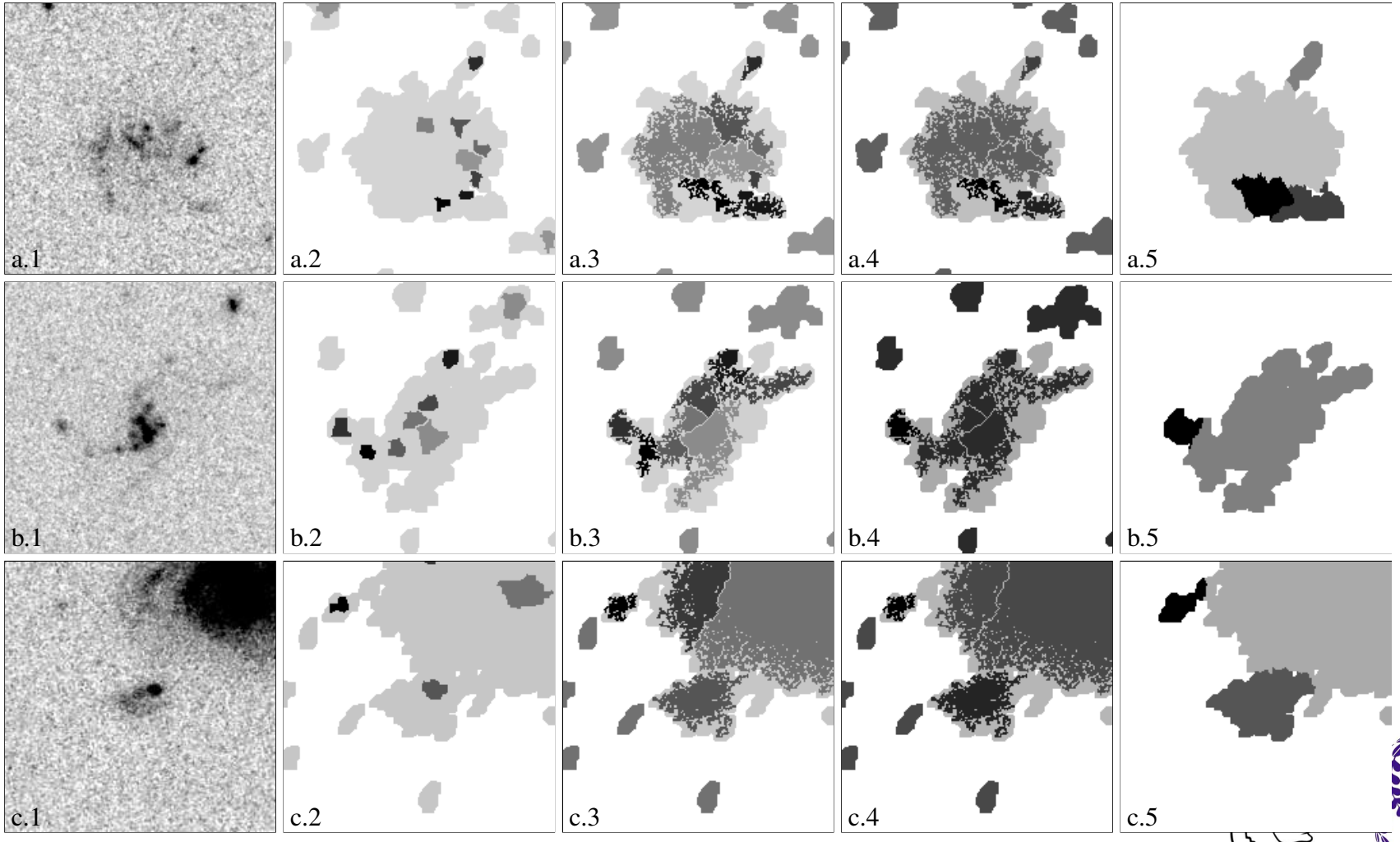


NoiseChisel – Segmentation – True clumps



clumps are found independent of user input.

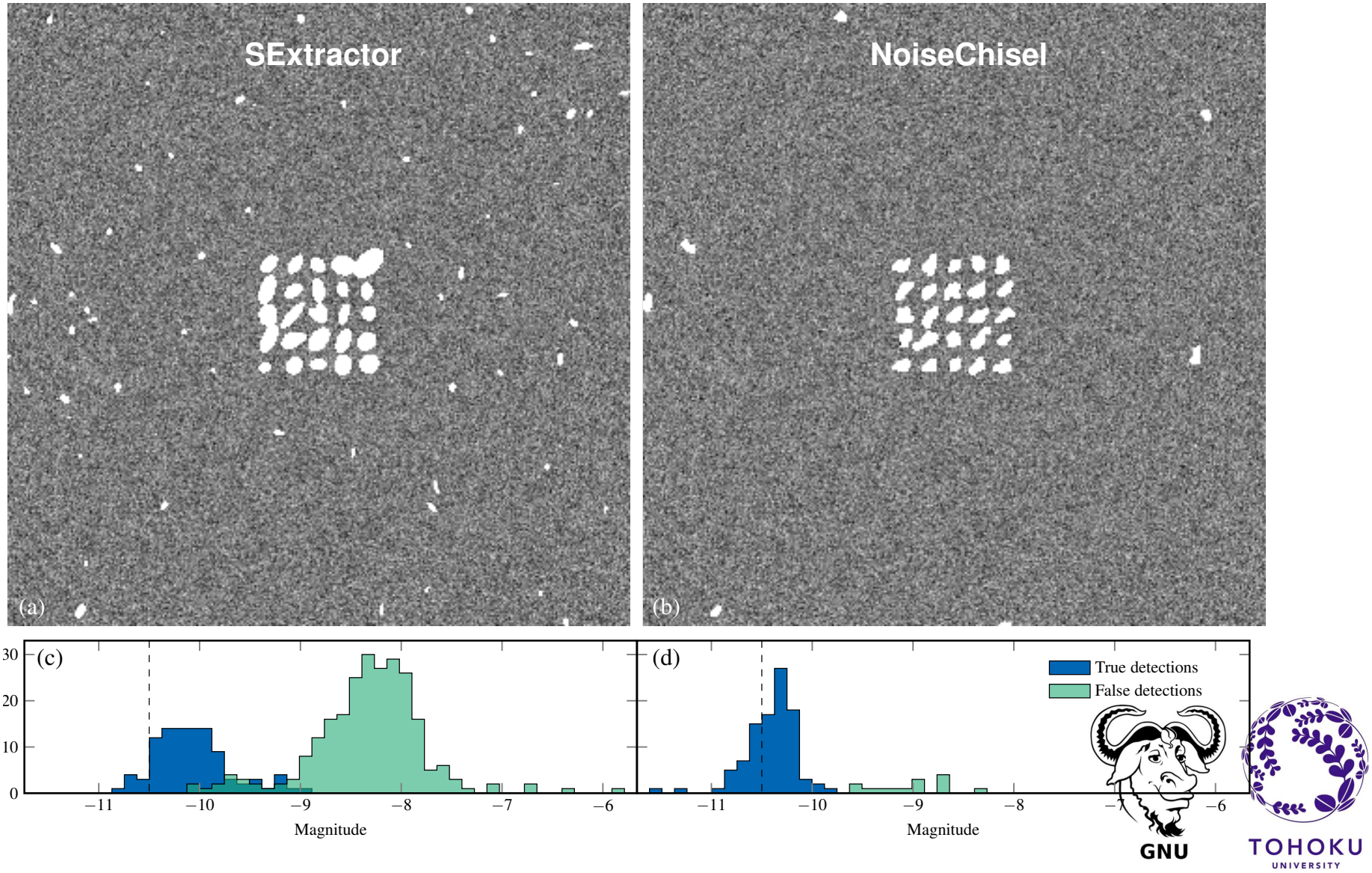
NoiseChisel – Segmentation – Objects



GNU

TOHOKU UNIVERSITY

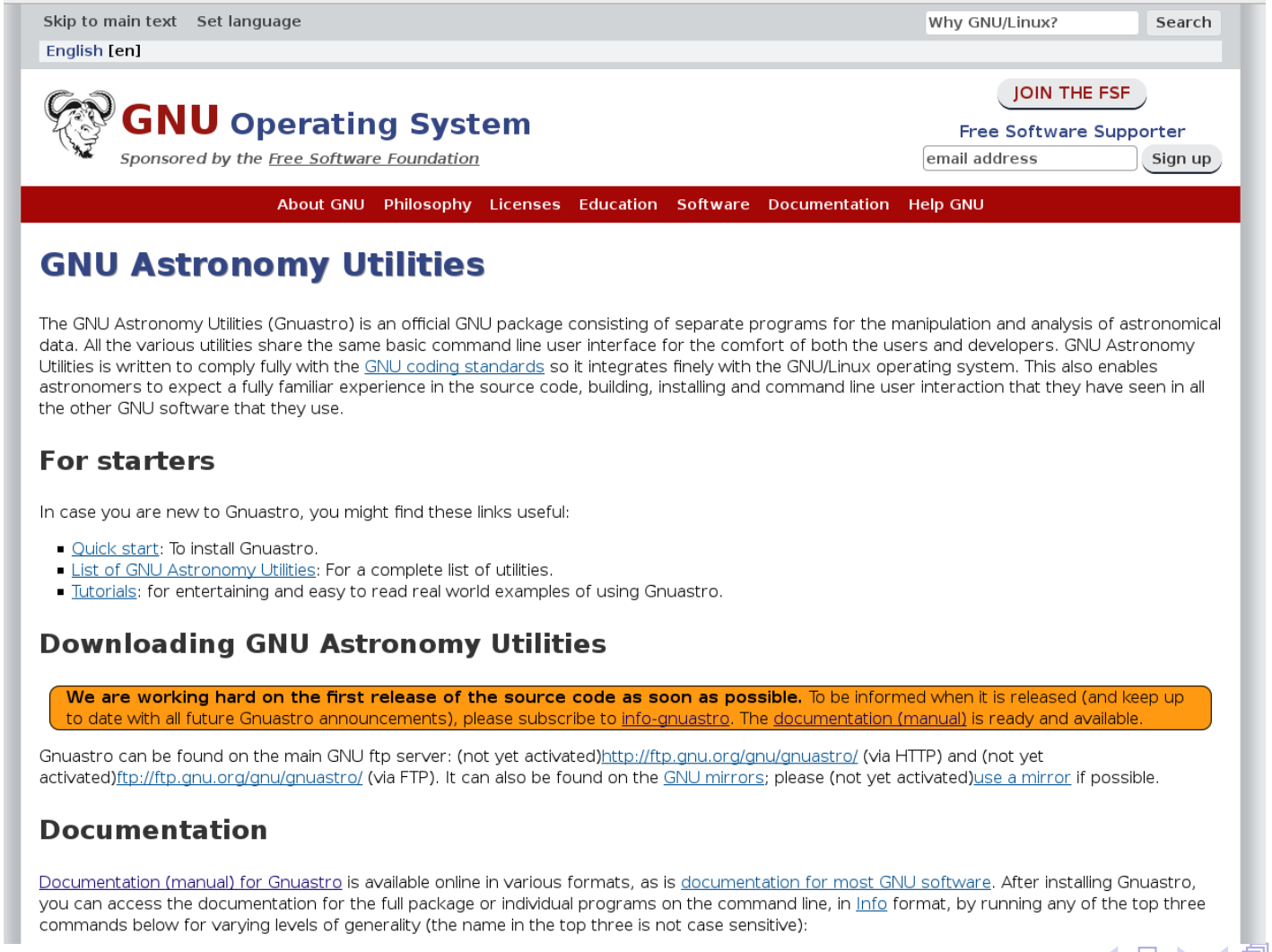
Purity and magnitude dispersion



GNU Astronomy Utilities (Gnuastro)

NoiseChisel is part of the GNU Astronomy Utilities

- https://www.gnu.org/software/gnuastro/



Current Utilities

- ConvertType
- Convolve
- Header
- ImageCrop
- ImageStatistics
- ImageWarp
- MakeCatalog
- MakeNoise
- MakeProfiles
- NoiseChisel
- SubtractSky

GNU

TOHOKU UNIVERSITY

Gnuastro's manual

GNU Astronomy Utilities

Astronomical data manipulation and analysis
for version 0.1, 7 October 2015

Mohammad Akhlaghi

Short Contents

GNU

TOHOKU UNIVERSITY

Gnuastro's manual

Table of Contents

GNU

TOHOKU UNIVERSITY

Gnuastro's manual

GNU

TOHOKU UNIVERSITY

Gnuastro's manual

give a negative, float or string value for a FITS image extension or column number. As another example, you might give a value larger than unity for an option that only accepts fractions (which are always less than unity and positive).

CAUTION: In specifying a file address, if you want to use the shell's tilde expansion (~) to specify your home directory, leave at least one space between the option name and your value. For example use `-o ~/test`, `--output ~/test` or `--output=~/test`. Calling them with `-o~/test` or `--output=~/test` will disable shell expansion.

CAUTION: If you forget to specify a value for an option which requires one, and that option is the last one, Gnuastro will warn you. But if it is in the middle of the command, it will take the text of the next option or argument as the value which can cause undefined behaviour.

NOTE: All counting in Gnuastro starts from 0 not 1. So for example the first FITS image extension or column in a table are noted by 0, not 1. This is the standard in C and all languages that are based on it (for example C++, Java and Python).

4.1.4 Common options

To facilitate the job of the users and developers, all the programs in Gnuastro share some basic command line options for the same operations where they are relevant. The list of options is provided below. It is noteworthy that these similar options are hard-wired into the programming of all of Gnuastro programs using GNU C Library's argument parser merging ability.

For some programs, some of the options, might be irrelevant for example MakeProfiles creates FITS images based on a given catalog. Therefore no input images (and thus HDUs) are necessary for it. In such cases, the option is still listed and if a value is given for it, it is completely ignored.

4.1.4.1 Input/Output options

These options are to do with the input and outputs of the various programs.

`-h`
`--hdu` (`=STR`) The number or name of the desired Header Data Unit or HDU in the input FITS image or images. A FITS file can store multiple HDUs or extensions, each with either an image or a table or nothing at all (only a header). Note that counting of the extensions starts from 0(zero), not 1(one). When specifying the name, case is not important so `IMAGE`, `image` or `ImAgE` are equivalent.

A `#` is appended to the string you specify for the HDU[1] and the result is put in square brackets and appended to the FITS file name before calling CFITSIO to read the contents of the HDU for all the programs in Gnuastro. CFITSIO has many capabilities to help you find the extension you want, far beyond the simple extension number and name. See CFITSIO manual's "HDU Location Specification" section for a very complete explanation with several examples.

[1] With the `#` character, CFITSIO will only read the desired HDU into your memory, not all the existing HDUs in the fits file.

`-o`
`--output` (`=STR`) The name of the output file or directory. With this option the automatic output names explained in Section 4.5 [Automatic output], page 43 are ignored.

`-D`
`--dontdelete`
By default, if the output file already exists, it will be silently replaced with the output of this run of all Gnuastro programs. By calling this option, if the output file already exists, the programs will warn you and abort.

`-K`
`--keepinputdir`
In automatic output names, don't remove the directory information of the input file names. As explained in Section 4.5 [Automatic output], page 43, if no output name is specified, then the output name will be made in the existing directory based on your input. If you call this option, the directory information of the input will be kept and the output will be in the same directory as the input. Note that his is only relevant if you are running the program from another directory!

4.1.4.2 Operating modes

Another group of options that are common to all the programs in Gnuastro are those to do with the general operation of the programs. The explanation for those that are not only limited to Gnuastro but can be called in all GNU programs start with (GNU option).

`--` (GNU option) Stop parsing the command line. This option can be useful in scripts or when using the shell history. Suppose you have a long list of options, and want to see if removing some of them (and using the default values) can give a better result. If the ones you want to remove are the last ones on the command line, you don't have to delete them, you can just add -- before them and if you don't get what you want, you can remove the -- and get the same initial result.

`--usage` (GNU option) Only print the options and arguments. This is very useful for when you know the what the options do, you have just forgot their names. See Section 4.6.1 [`--usage`], page 44.

`-?`
`--help` (GNU option) Print all options and an explanation. Adding this option will print all the options in their short and long formats, also displaying which ones need a value if they are called (with an `=` after the long format). A short explanation is also given for what the option is for. The program will quit immediately after the message is printed and will not do any form of processing. See Section 4.6.2 [`--help`], page 44.

GNU TOHOKU UNIVERSITY

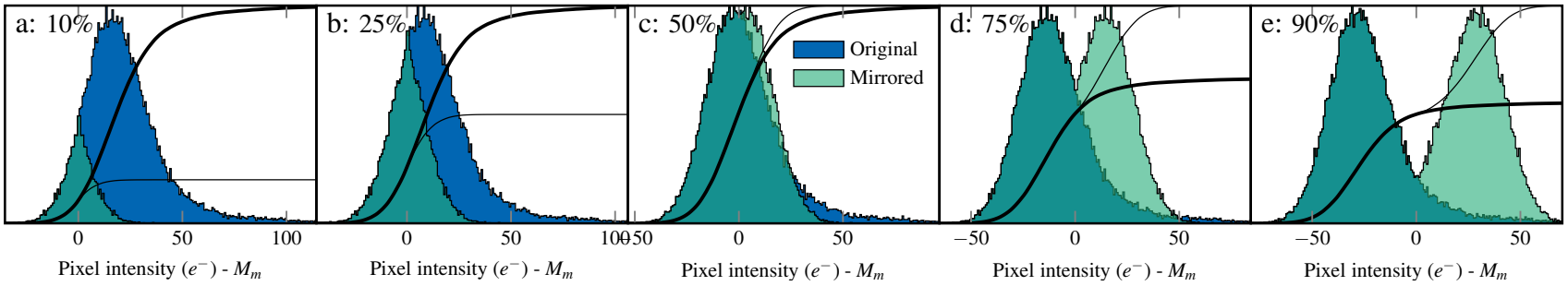
Reproducible science

See source of arXiv:1505.01664 (this paper) and/or `https://gitlab.com/makhlaghi/NoiseChisel-paper`.

- All the scripts and configuration files submitted to arXiv.
- Operation is managed by a Makefile.
- Data-generated numbers are LaTeX variables generated by the scripts.
- Input for all the plots are also generated by the scripts.
- Readers are encouraged to check/modify the parameters to see their effect.
- Everything fully explained in `reproduce/README`.

Finding the mode

- Data shifts the mode to the positive.
- The quantile is only comparable when the mode is.
- To find the mode, we define a "mirrored distribution" (green):



- The mode is where the cumulative frequency plot of the mirror distribution and original distribution are most similar: