

# Tim Baker lab software modification history (auto3dem, robem, and utility programs)

This document describes incremental modifications made to the Baker lab software suite. Since the image preprocessing, image reconstruction, format conversion, and utility programs are so tightly coupled (e.g. shared standards and libraries) and released as a single package, all modifications are described together in a single document. Owing to the complexity of the software, it is impractical to provide a line-by-line listing of alterations. Instead we focus primarily on modifications that impact the end user and on large-scale changes to the code. Starting on version 4.02, more details on some of the listed changes can be found in the file MINOR\_RELEASE\_NOTES, included in the software package.

## Table of Contents

<b>auto3dem v4.03 .....</b>	<b>2</b>
<b>auto3dem v4.02 .....</b>	<b>4</b>
<b>auto3dem v4.01 .....</b>	<b>7</b>
<b>auto3dem v4.0.....</b>	<b>8</b>
<b>auto3dem v3.15 .....</b>	<b>11</b>
<b>auto3dem v3.14 .....</b>	<b>12</b>
<b>auto3dem v3.13 .....</b>	<b>13</b>
<b>auto3dem v3.12 .....</b>	<b>15</b>
<b>auto3dem v3.11 .....</b>	<b>16</b>
<b>auto3dem v3.10 .....</b>	<b>17</b>
<b>auto3dem v3.09 .....</b>	<b>18</b>
<b>auto3dem v3.08 .....</b>	<b>19</b>
<b>auto3dem v3.07 .....</b>	<b>19</b>
<b>auto3dem v3.06 .....</b>	<b>20</b>
<b>auto3dem v3.05 .....</b>	<b>21</b>
<b>auto3dem v3.04 .....</b>	<b>21</b>
<b>auto3dem v3.03 .....</b>	<b>23</b>
<b>auto3dem v3.02 .....</b>	<b>23</b>
<b>auto3dem v3.01 .....</b>	<b>24</b>

<b>auto3dem v3.0</b> .....	<b>24</b>
<b>auto3dem v2.04</b> .....	<b>26</b>
<b>auto3dem v2.03</b> .....	<b>26</b>
<b>auto3dem v2.02</b> .....	<b>29</b>
<b>auto3dem v2.01</b> .....	<b>33</b>
<b>auto3dem v2.0</b> .....	<b>34</b>
<b>auto3dem v1.12</b> .....	<b>37</b>
<b>auto3dem v1.11</b> .....	<b>38</b>
<b>auto3dem v1.10</b> .....	<b>39</b>
<b>auto3dem v1.09</b> .....	<b>40</b>
<b>auto3dem v1.08</b> .....	<b>42</b>
<b>auto3dem v1.07</b> .....	<b>43</b>
<b>auto3dem v1.06</b> .....	<b>44</b>
<b>auto3dem v1.05</b> .....	<b>44</b>
<b>auto3dem v1.04</b> .....	<b>45</b>
<b>auto3dem v1.03</b> .....	<b>46</b>
<b>auto3dem v1.02</b> .....	<b>46</b>
<b>auto3dem v1.01</b> .....	<b>47</b>

## **auto3dem v4.03**

### **Compilation**

- Makefiles reorganized for better control of warnings
- Added make.inc.compflags where all flags are combined after being modified by single programs, if needed

### **General modifications**

- Extended limits on number of micrographs and parameter files
- Extended length limit for names of mrc files
- Almost all the files have been analyzed and modified in order to minimize the warnings at compilation

### **Common libraries**

- Fixed alignment of derived type pifHead. Problem seems specific to Intel compiler

- Changed many integer variables from 2-bytes to 4-bytes format

### **robem**

- Fixed linear gradient correction when applied to data with negative values
- Fixed blemish removal bug
- Removed excessive output messages to terminal
- Avoided generation of cmap.robem file
- Changed default values for thinning and dilation in boxing gui and inverted their sequence of operation: first thinning and then dilation.
- Improved autobox routine by modifying some preprocessing steps and adding a post-selection screening based on screening tails of score histogram.
- Added standard deviation to output of spherical average when invoked by command line
- Fixed bugs when handling maps larger than 1024 pixels
- Some subroutines moved out of robem.c, including statistical analysis of images and autoboxing

### **autopp**

- Added dilation parameter for autobox routine, option M
- Added parameter in option 4 to skip generation of dat files from ctf info if values are beyond given range
- Modified option 3 to delete temporary files generated for ctf estimation

### **auto3dem**

- Added *generate\_map* parameter to optionally split reconstruction and alignment into two separate steps. Accepted values are *only* (only reconstruction), *yes* (alignment and reconstruction), and *no* (only alignment)
- Added *partrad* parameter, which is dynamically updated to reflect the particle radius as estimated from the latest reconstruction
- Added *adapt\_angle* flag to determine adaptively the angle step for search/refinement by ppft/po2r, using the particle radius and the current resolution level of the reconstruction
- Changed criterion to determine when refinement supersedes search mode. Now the switch happens when the improvement between last two iterations is less than 0.5 angstroms
- Changed default parameters for ctf amplitude correction in po2r and p3dr in order to reduce the contribution of low frequency components
- Changed definition of capsid score, used to select the best candidate map during the random model computation

### **P3DR**

- Fixed determination of zero padding factor when requesting generation of smaller maps
- Fixed several routines to handle large maps (linear size  $\sim > 1300$  pixels)

- Memory footprint optimized, especially for serial version
- Added optional precompilation flag, in Makefile, for calculating the map in single precision. Useful to reduce the amount of memory needed

### **POR**

- Added substep refinement iteration around best orientation during local refinement
- Changed default weighting in the distance metric, which now accounts for the radial frequency

### **PSF**

- Changed calculation of frequency bins in order to be equal size rather than equal number of contributing voxels
- Modified FSC calculation to properly account for phase differences

### **tiff2pif**

- Fixed binning option for micrographs with short integer data format

## **auto3dem v4.02**

### **Compilation**

- Added compiler flag for large files support (4.01.12, 4.02)
- Updated flags for Intel compilers (4.01.3, 4.01.11)
- Minor cleanup and updates (4.01.8)

### **General modifications**

- Minor cleanup throughout code base and update to F90-style (4.01.9)
- Modified all I/O routines in order to access to large files (> 2GB) (4.01.12)

### **Common libraries**

- Code cleanup (4.01.9)

### **robem**

- Code cleanup (4.01.6, 4.01.8, 4.01.9, 4.01.11, 4.02)
- Minor reorganization of source files (4.01.1, 4.01.9)
- Minor fixes for improving portability (4.01.2)
- Fixed bug in rotational filtering (4.01.6)
- Box coordinates are saved in data headers of boxed images. This makes it possible to delete particle images and generate new bcrd files that are consistent with the image file (4.01.5)
- Added simple approach in autoboxing routine, to filter X-ray-corrupted pixels from micrographs (4.02)
- Statistics for removing blemishes from boxed particles is now calculated on enclosed circle rather than on the full box (4.02)

### **autopp**

- Fixed some typos (4.01.3)
- Modified options B (creating boxed image files) and K (recentering boxes, creating new bcrd and dat files) to better handle instances of multiple matches (4.01.5)
- Modified option E (batch conversion from tiff to pif format) to handle inversion of contrast (4.01.7)
- Modified option 2 (generating tiles from micrographs) to get directly tiles from micrographs instead of just their coordinates (4.02)
- Added option X to perform conversion from dm3 to pif format in batch mode (4.02)
- Added threshold argument to option M (autoboxing) for discarding low quality particles (4.02)
- Added argument to option 3 (ctf estimate) to output power spectrum from defocus estimation

### **auto3dem**

- Exposed several refinement modes available in POR. The modes are selected by the new field 'po2r mode', and they can be: local (original local refinement), global(search over a grid extending the full asymmetric unit, requires new parameter glob\_anglestep), mag (magnification refinement, requires new parameters nmagref, magref\_step, magref\_reset and magref\_calibrate), ticos\_equiv (search over the 60 symmetry-equivalent solutions). The change to interface is backward compatible with input parameter files from previous versions of auto3dem (4.02)
- Shell scripts made compatible with dash (default shell in Ubuntu flavor) (4.02)
- Minor bug fixes (4.01.11)

### **capsid\_generator**

- New program to generate PDB files listing coordinates of the points in an icosahedral lattice (4.01.10)

### **ctffiind3/ctftiltt**

- Code refactorization (4.01.9)
- Fixed bug in ctffiind3.f that was making the program hung because no tiles were selected (already fixed in the current Grigorieff's implementation).

### **ctffit**

- Program removed. Deprecated in favor of ctffiind3 (4.01.8)

### **dm32pif**

- New program to convert micrographs from dm3 format to pif. A binning option is included (4.02)

**emprj**

- Fixed bug on CTF parameters (4.01.11)

**imgstatst**

- fixed nearest neighbor interpolation used for calculating radial profile (4.02)
- Fixed typos (4.01.11)

**leginon\_to\_bcrd**

- Fixed bug in setting the box radius (4.01.9)

**mrc2pif**

- Improved ability to convert mrc files. Only 3D maps and 2D stacks of particles are automatically forced to odd size (4.02)
- Box radius written in the output file header (4.01.11)

**P3DR**

- Fixed bug in routine realtocomplx (4.01.4)

**PCTFR**

- Fixed bug that caused crashes with particles not well centered (4.02)

**pifmapmask**

- Minor updates (4.01.9)

**POR**

- Optimized the magnification refinement (po2r mode mag in auto3dem). Computational time for refinement of magnification is now comparable (or smaller) to an iteration of local orientation refinement. New features: - refinement around a previous solution (through option magref\_reset in auto3dem); - minimization of scaling of the output map, to use if the magnification is not intended for calibrating the data against a gold-standard (through the option magref\_calibrate in auto3dem) (4.02)
- Improved convergence of global search. Reference score for current solution is now calculated with latest settings, to guarantee that resulting map is an actual refinement of the previous one. Also, if particle orientation is changed, a local refinement follows.
- Fixed bug in magnification refinement mode, so that figure of merit stored in parameter file is correct (4.02).

**PPFT**

- Fixed bug to better estimate particle origins under all conditions. The estimate was always off by one pixel when using bin=2 and the original box size was such that the center was an odd number (e.g. box size: 215 -> origin: 107) (4.01.12)

- Added tests that avoid the program to interrupt without messages (4.02)
- Minor bug fix (4.01.12)

### **tiff2piff**

- Added 4th argument to command line to invert contrast (4.01.7)
- Code cleanup in I/O routines (4.01.9)

## **auto3dem v4.01**

### **autopp**

- Complete overhaul of program to make use more intuitive
- All output files generated in directory where autopp is run
- Input parameters can be provided at the prompts or specified on the autopp command line. If partial input is given, user will be prompted for remaining arguments.
- autopp command line written out if input provided at the prompt
- Option 4 (generation of particle parameter files from ctffind3 output) generates warnings if defocus levels or degree of astigmatism falls outside of user specified bound.
- Particle reboxing greatly simplified and provides two options for placement of new box coordinate and particle parameter files.

### **PPFT**

- Bug fixes implemented so that program does not crash if unable to determine particle origins. Results are still meaningless for these particles, but segmentation faults are avoided.

### **robem**

- Bug fix so that defocus major and minor values are not assigned values before pixel size has been set
- Radial in radial projection window now shown in both pixels and Angstroms

### **ctffind3**

- Code optimized to run up to 3-3.5x faster than previous version under certain conditions (typically smaller micrographs, larger defocus search range, and smaller defocus search step size)

### **pifmapmask**

- Replacement for pifmapdel that provides options both for deleting or masking volume covered by PDB structure.

### **parse\_dm3**

- Note added to output alerting users that certain hardcoded parameters are specific to Baker lab microscope calibration

## **fixpif**

- General cleanup to simplify code and improve maintainability

## **auto3dem v4.0**

**NOTE** – Following change had been implemented in v3.15, but was not described in previous version of this document:

Simplified the process of selecting CTF nodes in robem “FFT – CTF Estimation” window. By default, images are assumed to be stigmatized, but user can toggle between choosing parameters for stigmatized or astigmatic images.

## **Building executables**

- The make\_all script unsets the LANG environment variable before building executables. In some cases the default value for LANG can interfere with the build, particularly for robem and other GUI applications.

## **auto3dem**

- Improvements made to program imgstats to provide additional statistics on either PIF maps or the individual images in a boxed image file. Program has also been integrated into auto3dem to write out map statistics after all iterations.
- No longer need to specify “ppft verbose -1” in the auto3dem master file in order to avoid expensive calculations of 3<sup>rd</sup> correlation coefficient (cmp\_cc). Program PPFT has been modified so that verbose = -1 is the default.
- Automatic determination of capsid inner and outer radii is now much more reliable and faster to compute. Radii are determined directly from radial density profiles and no longer rely on the cmp\_cc coefficient from PPFT.
- Parallel jobs can now be launched by auto3dem using either mpirun or aprun, thereby providing greater flexibility regarding parallel computing environment. At run time auto3dem checks for the existence of these programs and automatically chooses the appropriate one.
- Negative values in FSC curve set to zero before attempting to apply noise suppression algorithm in P3DR. Avoids generation of pathological map where all pixels have the same value.
- Source code (include/allprogs.inc) updated so that maximum number of CPUs is now 512.

## **Defocus estimation using ctffind3**

- The ctffind3 and ctftilt programs (Grigorieff lab) have been integrated into the auto3dem software. The source code is located in the ctfniko directory



and the executables are built at the same time as other auto3dem applications when the make\_all script is run. These executables have been named ctftilt\_a3dm\_mp and ctffind3\_a3dm\_mp to avoid conflicts with existing installations. Defocus estimation can be done on multiple PIF or MRC micrographs using autopp Option 3 and particle parameter files generated using Option 4.

**NOTE** – ctffind3 and ctftilt use a different convention for the astigmatism angle than the Baker lab software. For use in auto3dem, the negative of the astigmatism angle should be used. This is done automatically when using autopp Option 4.

### **setup\_rmc**

- Capsid radii are estimated from the density profiles of the maps and automatically updated after all iterations. This greatly improves the likelihood that an individual random model computation (RMC) will converge to the correct low-resolution structure.
- The auto3dem master file generated by setup\_rmc sets the default number of iterations to 10, no longer generates the commented out “ppft verbose -1” line, and comments out the “auto bin\_reduce 1” line. The latter change was made since running PPFT with un-binned data before transitioning to auto3dem refine mode typically does not lead to higher resolutions.
- The algorithm used to select the best model from multiple models has been made more reliable. It was found that FSC-based methods did not always choose the best model, particularly for smooth, featureless capsids.
- The original RMC can be selected with the -trad flag.

### **robem**

- Major overhaul of code to avoid use of User Interface Language (UIL) files for the specification of initial state of GUI. This greatly improves code readability, portability, and maintainability.
- Isolation of many GUI operations so that robem can be run in batch mode (normally through control of autopp) without having to launch Xvfb. Before making these changes, getting robem to run in batch mode was very problematic and often the biggest hurdle to portability. No changes are necessary when using autopp since modifications to robem syntax have been implemented. When running robem standalone in batch mode, need to make following change to calling sequence

OLD: xbatch robem -pre infile [other options]

NEW: robem -nogui -pre infile [other options]

- Fixed the feature that adds Gaussian noise to an image. This had been done incorrectly in previous versions as evidenced by the fact that adding

Gaussian noise with mean=0 changed the mean intensity of the image.  
(This feature is found at Manipulate > Fix Values)

- Fixed bug that caused robem to crash when clicking in drawing area of particle processing screen before particles had been loaded.
- Fixed bug where display of radial projections would sometimes cause robem to crash.
- Modified code so that error messages no longer generated when performing zooming on a micrograph.
- Overhauled "Point Value" screen so that data is easier to interpret, distances are provided in Angstroms if the pixel size is listed in the map/image header, and behavior of mouse buttons is consistent with behavior that most users would expect.

### **gensph**

- Program for generating a spherically averaged PIF map from an input map

### **pifmapdel**

- Program for masking a PIF map using PDB file coordinates

### **prjcompare**

- Program for calculating the difference between two image stacks. Typically used to compare boxed images to projections of a map.

### **parse\_dm3.pl**

- Script for parsing DM3 file header and extracting voltage, spherical aberration, magnification, and pixel size. The latter quantity is specific to the Baker lab and the get\_pixel\_size function will have to be modified for other microscopes.

### **PIF library and calls to PIF library**

- Initialize the character string used to specify the read/write arguments that are passed to pif\_open. On some machines, approximately 1-in-1000 calls to pif\_open from P3DR randomly set the second character in the string to a value that caused the program to crash.
- General cleanup of PIF library, including standardization of commenting and elimination of obsolete code related to storage of binned data on disk.
- Obsolete functions or functions with redundant capabilities were eliminated. Applications that use the PIF library were modified accordingly. The list below contains the deleted functions, with the replacement calls, if needed, in parenthesis:
  - 1) pif\_read\_tmap2 (pif\_read\_map2)
  - 2) pif\_read\_img2 (pif\_read\_short\_image)
  - 3) pif\_read\_img1
  - 4) pif\_read\_byteshort\_image

## auto3dem v3.15

### auto3dem

- Bug fix to P3DR so that it can handle case where the first particle parameter file is empty. Behavior when subsequent parameter files are empty is unchanged.
- Modifications to setup\_rmc to write scripts that simultaneously launch and manage multiple random model calculations under control of Portable Batch Scheduler (PBS). All currently running jobs are terminated and scheduled jobs de-queued if any calculation terminates early upon finding a suitable starting model.
- Add new keywords to restrict particles used in reconstruction on the basis of orientation angle omega. Select all values within omega1\_tol of omega1 OR omega2\_tol of omega2, while properly accounting for behavior at 0°/360°. By default, all values of omega are selected.

# Example 1 – select omega within 10° of 0°

```
auto omega1      0
auto omega1_tol  10
```

# Example 2 – select omega within 20° of 90° of 270°

```
auto omega1      90
auto omega1_tol  20
auto omega2      270
auto omega2_tol  20
```

### robem

- New buttons added to FFT dialogue so that user can choose between writing full (header + particle records) or empty (header only) particle parameter files.
- When writing out boxed image and box coordinate files, default names automatically provided. For example, if the name of the micrograph being processed was micro.pif, the default image and coordinate files would be micro\_box.pif and micro.bcrd, respectively.

### imgstats

- New program for calculating statistics on maps and boxed image stacks. Provides global statistics and radial profile information. Note that output for image stacks can be quite large since a radial profile will be generated for each image.

### Code cleanup

- In addition to modifications listed above, numerous changes were made to the code to improve portability and maintainability. Debug code and old code that had been retained for compatibility with pre-ANSI C was removed.

Unused functions were deleted and redundant routines were merged into single versions. Obsolete or unnecessary variables were removed from argument lists and required changes to calling routines were made. A number of files that were no longer needed after software cleanup were removed: symlib.c, imsubs2.f, parser.f, libR/appodize.c, emicolib.f, libCtf.c, vax\_minmax.inc, em.cmm, pftlib.f, movc3.f.

- **Note to developers:** Major cleanup of the PIF library was initiated. More than 2000 lines of code were eliminated and redundant routines with very similar functionality were merged. All codes distributed with v3.15 have been modified to use the new PIF library, but old applications will either need to continue using the earlier version or be changed to handle the new interface.

## auto3dem v3.14

### auto3dem

- Feature added to POR to test the 60 orientations related to the starting orientation by icosahedral symmetry. Useful when trying to identify lower symmetry components of an otherwise icosahedral structure. Enabled as shown below

```
por ticos_equiv 1    # Add to auto3dem master file
ticos_equiv 1       # Add to POR input file
```

- Feature added to POR to perform global search of orientation space. Useful when trying to ensure that all image orientations have settled into the correct part of the asymmetric unit. Use with caution since this option can VERY expensive since it involves a 3D search over  $(\theta, \phi, \omega)$ . Enabled as shown below

```
por global_por 1    # Add to auto3dem master file
global_por 1       # Add to POR input file
```

- Added tests to POR and PPFT to make sure that input map is present and readable. Program terminates and error message written if readable map not found.
- P3DR now writes out the number of empty voxels at each spatial frequency. This output can be used as a diagnostic to determine if transform of map is being too sparsely populated at a particular spatial frequency.

### robem

- General improvements to robem including:
  - Setting of sensible defaults (e.g. 1D defocus estimation rather than 2D)

- More meaningful labels provided
  - Widgets resized and repositioned to fit properly on screen
  - Clearer behavior of toggle buttons
- Code modified so that bulged icosahedral projections are displayed properly. For large maps, would sometimes see artifacts at equator.
- Partial overhaul of code to improve readability and maintainability.

### **oned, ctfdisp, fixpif**

- No longer need to set EMDIR environment variable in order for programs to find corresponding UID configuration file. File now found by either searching PATH variable or determining location where program is installed when full path is specified to executable. General cleanup of code and elimination of unused routines.

### **empri**

- Eliminated discrepancy between code and documentation for CTF parameter list. Cc, beta, and dE must now be provided.
- Fixed bug so that projections of map are now multiplied by CTF. Previously empri would fail to apply any correction when one of the “multiply by CTF” options was chosen.

### **New PDB manipulation scripts**

- Although these are not strictly related to image reconstruction or analysis, they may be useful for performing basic tasks related to constructing starting models from PDB coordinates.

pdbcalpha – retain only C $\alpha$  atoms.  
 pdbclean – retain only TER, REMARK, and ATOM records  
 pdbcomb – combine two PDB files  
 pdbscale – scale atom coordinates

### **New box file format conversion scripts**

- The EMAN boxer program and robem produce box coordinate files in different formats. To enable easy conversion between the two, the following scripts have been created

box2bcrd – convert EMAN box format to robem bcrd format  
 bcrd2box – convert robem bcrd format to EMAN box format

## **auto3dem v3.13**

**NOTE** – As of version 3.13, when the software suite is built using the make\_all script, the symbolic link BIN pointing to the bin directory is no longer created. In your configuration file, make sure that your path is set appropriately.

Also (see robem below), the EMDIR environment variable no longer needs to be set. It is now sufficient just to set the path to point to the location of the executables.

### **auto3dem**

- Backup of summary file automatically created if launching a new, as opposed to a continuation, run. Avoids accidental overwrite of results summary.
- Format of summary file modified to include number of micrographs and range of defocus values used in the reconstruction.
- More robust error detection and reporting. Error messages provided showing parallel program and input file being used when abnormal termination occurred.
- Random model computations modified to handle concatenated particle parameter files and image stacks in MRC format. When launching setup\_rmc, data directory automatically searched for parameter files with suffixes as large as 999.
- Output routine called by P3DR modified to avoid MPI buffer overflows and allow writing of maps of size larger than 10243.
- Cleaner handling of symmetries other than icosahedral (532). Specification of symmetry in auto3dem master file using “auto symm\_code n” overrides symmetry parameters set for all other programs.
- Control of handedness tests in POR now under control of input parameter. Can be set either in auto3dem master file using “por handtest [0|1]” or POR input file with “HAND [0|1]”. By default handedness test is still performed.
- Magnification factor refinement capabilities added to POR, but not yet integrated into auto3dem. Controlled using “MAGR nmagf magfstep” in POR input file, where nmagf is number of steps to take in each direction from magnification factor equal to 1.0 and magfstep is the step size.
- Particle images with centers that are too close to edges of box (< 10% box width) are excluded from origin and orientation refinement in POR. This is done to prevent unexpected termination of program.
- Flush operations added to P3DR to force writes of buffered data. This makes it easier to track progress of reconstruction and identify input images that cause P3DR to crash.

### **robem**

- No longer need to set EMDIR environment variable in order for robem to find corresponding UID configuration file. File now found by either searching PATH variable or determining location of robem when full path is specified to executable.

- Difference map now printed along with “second” map from robem difference map window.
- Modification to “distance” feature in Point window so that distances can be measured when working with image stacks.

### **ctfdisp**

- Display fixed so that resolution and value of CTF can be easily read.

In addition to changes listed above, extensive cleanup of underlying code: removal of obsolete functions; deletion of unused variables, common blocks, and macros; addition of INTENT specifications in Fortran subroutines; placement of frequently used sets of operations into functions; general cleanup to improve readability of code; simplification of I/O for maps and image data; modifications to enforce compatibility of C header and Fortran include files.

## **auto3dem v3.12**

### **auto3dem**

- Automation script and all programs that read image data (PO2R, P3DR, PPFT, PCTFR) modified to handle both MRC and PIF image stacks. Note that the input maps must still be in the PIF format and all output will also be generated in PIF format. This is the first step towards making the software completely compatible with MRC data.
- Particle images more than 50% of the box radius from center of box automatically rejected from reconstruction. Default can be overridden by specifying a value for the “auto\_box\_center\_offset” parameter.
- Files listing the images that had been rejected from the reconstruction (e.g. as a consequence of particle origin being too far from center of box) are now generated and stored in the project’s dat/virus\_INTFILES directory.
- Programs PO2R, P3DR, PPFT, and PCTFR now detect and automatically reject particle images that have invalid image numbers (negative, zero, or greater than the number of images in the corresponding image files). The presence of invalid image numbers formerly caused these programs to crash, but now simply results in a warning message in the output file.
- Image files for which the corresponding particle parameter files (.dat files) are empty are now properly handled by program P3DR. Although these micrographs never contributed to the reconstruction, they could occasionally cause P3DR to crash. Note that if ALL parameter files are empty, it is still assumed that ALL image data will be used.
- Tests on the microscope accelerating voltage and amplitude fraction made less stringent. For reconstructions from negative stain image data, the amplitude contrast is generally much larger than that for cryo and valid data sets were sometimes rejected by auto3dem.

- PO2R modified to handle image magnifications greater than one. In previous versions of code, higher image magnifications could occasionally cause segmentation faults.
- Summary file now lists the total number of particle images that could have potentially been used in the reconstruction.

#### **autopp**

- Particle reboxing (option K) now allows the user to specify the current working directory as the location of the particle parameter (dat) and/or coordinate (bcrd) files. New files would be written to the dat\_new and/or bcrd\_new directories.

#### **setup\_rmc**

- Command line options added to control number of bins and resolution ranges used when calculating FSC curve.

#### **robem**

- Cursor used for boxing particles has been changed from “cross” to “diamond cross” shape, making it easier to see and track the cursor.
- Modifications made to repair bug that affected deletion of images from file.
- Text color changed on CTF estimation page to improve readability.

In addition to changes listed above, general cleanup of code base and removal of obsolete or redundant subprograms. make\_all and clean\_all scripts modified to use /bin/bash rather than /bin/sh since some flavors of Linux (e.g. Ubuntu) no longer support /bin/sh.

## **auto3dem v3.11**

#### **auto3dem**

- Program P3DR modified to handle maps of size larger than  $\sim 1200^3$ . Similar modifications made to PO2R, PCTFR, and PSF to allow calculations on larger maps. Map size limit in the latter set of programs was estimated to be  $\sim 1700^3$ .

#### **autopp**

- Improvements made to option K (generation of new box coordinates from particle origins and original box coordinates) to provide greater flexibility in naming convention of bcrd files relative to dat files.

#### **robem**

- Bug fix so that “Find mag” feature in difference map window works correctly.



- Modifications to automatically shift boxes that overlap edge of micrograph to lie entirely within the micrograph. Also made minor changes to code so that boxes very close to edges and corners of micrograph are handled correctly.

### **combine\_fsc**

- New program that combines the results from multiple Xmgrace .agr files to create a plot containing multiple FSC curves

### **Code cleanup**

- Eliminate unused functions, files, and variables
- Remove redundant files
- Simplify logic of setting labels, titles, and values in robem widgets
- Replace non-standard function calls to float() with Fortran 90 real( )

## **auto3dem v3.10**

### **auto3dem**

- Produce FSC curves using Xmgrace instead of the GD package. PNG file is generated only if Xmgrace is found in the Linux/UNIX path, but project file is always written to allow manipulation or plotting of FSC curves after completion of calculations. Using Xmgrace rather than GD circumvents the known problem that GD has in plotting data with numerical (non-evenly spaced) x-values. Modifications were made to the auto3dem.pl file and the new modules/grace\_utils.pm file was added.
- Modify logic so that FSC curve is only calculated if the 'auto estimate\_res' option is set and move the creation of the filtered files inside the block of code used to create full map.
- Fix bug in libCommpk/ctf\_para.f so that astigmatism is treated correctly. Had been making call to atan2(x,y) instead of atan2(y,x).
- Modify pftsearch.F and global\_cc.F so that we either calculate projections of model using the origins listed in the .dat files or, in the case where all .dat files contain only header information, the center of the image box. Get rid of PPFT files map\_sym\_cavg\_ppft.f and prjavg\_fft.f since these are no longer needed and update Makefile accordingly.
- Modify algorithm used in PPFT to minimize memory usage during generation of projections of the model. Required addition of new subroutine map\_prj\_slice.

### **autopp**

- Option F (blemish and linear gradient removal, normalization) modified to test that all boxed image files in the argument list are readable.
- Option 3 added to allow batch processing of micrographs with ctfilt.
- per\_ptle\_ctf
- Minor bug fix so that -mode astig works properly

### **robem**

- Modification and cleanup of writeBoxCoords functions so that bcrd files are written out immediately rather than waiting until completion of session.
- Default particle parameter files now have .dat\_000 rather than .dat\_001 suffix. This change makes it easier to correlate parameter files with auto3dem iteration numbers.

## **auto3dem v3.09**

### **auto3dem**

- Number of CPUs can now be specified using both -ncpu and -np flags
- Tests added to ensure that the inner and outer radii specified for programs PCUT and PPFT are both properly ordered (e.g. inner radius < outer\_radius) and less than radius of boxed images
- max\_cpu flag parameter now written to auto3dem restart and continue files for programs PO2R, P3DR, PPFT, and PCTFR
- Bug fix in PPFT to allow computations on maps larger than 8003
- Bug fix in PO2R to prevent out-of-bounds array access when using particle magnifications that are greater than 1.0

### **setup\_rmc**

- New command line option to specify maximum resolution of computed map
- New command line option to set radii used by programs PPFT and PCUT

### **autopp**

- Generation of new boxed coordinate (bcrd) files from particle parameter files and old bcrd files now more robust with regards to file formats and naming conventions
- Bug fix to option G (auto-boxing, blemish and linear gradient removal, normalization) to avoid entering infinite loop

### **per\_ptle\_ctf**

- Added capability to handle missing or out-of-order records in particle parameter files

### **robem and ctfdisp**

- Corrected spherical aberration of objective lens for FEI Polara and Sphera microscopes from 2.0 mm to 2.3 mm and 2.26 mm, respectively. Previously solved structures should still be valid since incorrect values had negligible impact on CTF correction.
- Modifications to allow calls to Fortran from C using name mangling conventions employed by PathScale compilers

- General cleanup of code to consolidate functions that handle behavior of arrow widgets
- Collection of macro definitions into single location to avoid accidental specification of conflicting definitions

## auto3dem v3.08

- I/O routines modified to handle embedded blank lines in particle parameter files
- Functions `parse_key_input( )` and `list_key_input( )` temporarily added back into `libUtil`. Needed in anticipation of possibly adding serial Fourier-Bessel image reconstruction program `EM3DR` to software release.
- Program terminates and error message written if attempting to use the following options together. Setting `PPFT verbose` flag to `-1` disables calculation of quantities needed to calculate inner and outer radii of annulus

```
auto freeze_annulus 0
ppft verbose -1
```

Made changes to the source code that allows all programs except `robem`, `ctfdisp`, and `emprj` to be built using the `PathScale` compilers.

- Variable types changed from `integer*2` to `integer*4` so that arguments to `mod()` intrinsic function are of the same type.
- Parenthesis added around negated terms so that expressions like `x = y + (-z)` are parsed correctly.
- Created double-underscored versions of functions in `PIF` library to allow for multiple name mangling conventions. For example, `func( )` → `func_( )` → `func__( )`.
- Explicitly added `#include <stdlib.h>` to C source files where needed.

## auto3dem v3.07

`AUTO3DEM` and all programs that are called by it (`P3DR`, `PCTFR`, `POR`, and `PPFT`) can now handle particle parameters files containing data for multiple boxed image files. These concatenated files are of the following form

```
image_file_name
defocus information
zero or more particle records
image_file_name
defocus information
zero or more particle records
```

...

No embedded blank lines are allowed in the concatenated particle parameter file, but trailing blank lines are still permitted.

Note that the script `setup_rmc.pl`, which is used to setup the random model calculations, still requires that each particle parameter file contain data for only a single micrograph. This may be remedied in future releases.

Other changes in this release include:

- All long lines (>72 characters) in Fortran source code have been shortened to fit within the standard fixed-format limit. This affected a small number of lines that were still syntactically correct after being truncated. The only effect of these lines was on the printing of rarely encountered error messages and in the display of the “signal minus background” curve in the RobEM CTF estimation screen.
- Warnings are no longer produced when defocus values are outside of the range 0.8-4.0 $\mu$
- Deleted the return statement in `pif2ccp4` main program. This had been causing errors for some newer versions of compilers.
- Added `xbatch.mscp` to `bin/` directory to allow RobEM to be run in the background in batch mode.
- Addition of new options to `autopp`: renaming of file prefixes and generation of BCRD file for tiling scanned image file.
- Make script now moves (instead of copying) all executables into `bin/` directory.
- General cleanup of code for programs `fixpif`, `ctfdisp`, and `oned`.

## auto3dem v3.06

- Modified image reconstruction programs to handle up to 2048 particle parameter files.
- Bug fix to programs `mrc2pif` and `pif2mrc` to allow proper conversion of image data in BYTE format.
- Bug fix to RobEM map projection feature in general and generation of projections spanning the icosahedral asymmetric unit in particular. Delete routines from RobEM source and replace with calls to `libCommpk`.
- Bug fix to 3D Display panel in RobEM that caused min/max radius data to be displayed in the wrong text box. This bug was introduced in v3.05 and had no effect on RobEM results.
- Increased width of text box that lists file names in Circular Average display. Previous box width was too narrow to allow longer file names to be read.

- Clicking on section button on main tool bar now displays central section of map rather than first section.
- Point screen panel now contains button for identifying and drawing marker at the center of a section.
- Improvements to handling of early termination mechanism in random model calculations
  - RMC\_DONE file automatically deleted during the random model calculations.
  - Existence of the RMC\_DONE file no longer triggers termination of auto3dem. This logic has been moved into the auto-generated RMC\_run script.
  - Message is written both to stdout and the auto3dem log file informing the user that the “auto quit\_early” flag had been set and that this feature is normally used only during the random model calculations.

## auto3dem v3.05

- Modified communications in P3DR (exch\_intp.F) and PPFT (pftsearch.F) to avoid MPI buffer overflows when working with large maps.
- All Fortran source code updated so that lines are less than or equal to 72 characters. Compiler flags used for extended Fortran source have been deleted from Makefiles and options added for PathScale compilers.
- Declaration of maximum number of particle parameter files and length of filenames now declared in a single include file (sizes.inc) that is used throughout all image reconstruction codes.
- Arguments for SYSTEM\_CLOCK function changed from integer\*8 to default integer.
- PIF library modified to write out parameters a0-a5 from CTF estimation procedure to global file header.
- Added capabilities for interpolation of packed boxed images to RobEM.
- Major cleanup of RobEM source code: improved formatting of loops and blocks; removal of non-informative comments; standardization of function layouts; removal of common macros/definitions in UIL and header files, placement into common definitions file, and modification of Makefile to dynamically generate content; deletion of unused or obsolete functions; combination of all widget deactivation functions into a single generic function.

## auto3dem v3.04

A number of important new features have been implemented in this release.

- CTF corrections can now be applied on a per-particle image basis in programs PPFT, PO2R, and P3DR. Three steps are needed to use this feature. Note that the program CTFTILT is not maintained by our laboratory and must be downloaded from [http://emlab.rose2.brandeis.edu/grigorieff/download\\_ctf.html](http://emlab.rose2.brandeis.edu/grigorieff/download_ctf.html)

1. Run CTFTILT program on micrographs.
2. Generate extended particle parameter files from CTFTILT output file and standard format particle parameter files using `per_ptle_ctf` script
3. Add the following line to the `auto3dem` input file, where `auto` enables the feature for all three programs

```
(auto|ppft|po2r|p3dr) per_ptle_ctf 1
```

- Noise suppression algorithm described in Rosenthal and Henderson, JMB 333 721-745 (2003) is now available in P3DR. Add following line to `auto3dem` input file

```
auto noise_suppression 1
```

- Program PPFT can now be run using all empty (except for header) particle parameter files. Number of particle images and box center are obtained from corresponding boxed image files specified in particle parameter files.
- Programs PPFT, PO2R, PCTFR, and P3DR now allow embedded comments in their input files. Everything after the first '#' character is ignored. Previous versions only allowed comments where '#' was the first non-whitespace character. Note that this feature only impacts users who run these programs standalone (not through `auto3dem`).
- Improved auto boxing of empty particles in RobEM, bug fix when writing movies with .pcx file naming, and enhancement of help facility.
- Variables specifying number of views per read deleted from all source code. No longer needed since overhaul of parallel algorithms in `auto3dem` v3.03. Corresponding parameters now ignored and marked as deprecated in `auto3dem`.

#### Other minor modifications and bug fixes in this release

- Cleanup of `#include` statements in RobEM source code
- Overhaul of `UIDATE` subroutine in `unix.f` to improve portability
- Rearrangement of order of elements in user defined type in `strucfac.inc` for better alignment of variables
- Explicit casting of variable in `tiff2pif.c` to avoid compiler warnings

- Fix error in specification of argument intents in pftcc\_peak.f
- Minor bug fix related to error handling during terminal input in sflib.f
- Change MIN0 and MAX0 function calls to MIN and MAX respectively in misclib.f
- Minor changes to make\_all and clean\_all
- Deletion of robem.uil – superseded by new\_robem.uil
- Deletion of libCompar/read\_files.F and libCommpk/readorient.f since they are no longer needed after completing cleanup of I/O
- Replacement of get\_ptle\_params.f with get\_ptle\_ioom.f and get\_ptle\_ioomctf.f to allow more flexibility in handling per-micrograph and per-particle CTF corrections
- Renamed map\_prj\_ppft.f to map\_prj.f and moved to libCommpk; deleted routines map\_prj, map\_prj\_axis, map\_prj\_xz, and map\_prj\_all from maplib.f; modified pftsearch.F to call map\_prj instead of map\_prj\_ppft; modified emprj.f and emprj\_xdy.f to use call new version of map\_prj.

## auto3dem v3.03

- Fixed bug that affected particle selection from PPFT-generated parameter files.
- Parameters added for new microscopes in robem.
- Minor modifications to avoid compiler errors from newer versions of gcc/gfortran.
- quick\_omega (PPFT) and quick\_search (PO2R) now turned on by default. These changes only impact users who run these as standalone programs since the options had been automatically enabled when running through AUTO3DEM.
- Improved commenting and elimination of unused code in robem, oned, and ctfdisp.
- Simplification of correlation coefficient calculations in PPFT. Changes do not affect code performance or results, but greatly improve readability and maintainability.

## auto3dem v3.02

- Overhauled parallelization scheme for programs PPFT, PO2R, and P3DR. Reduced memory requirements in PPFT by approximately 40%. Test problems run on 16 processors showing reductions in run time of 25-40% relative to v3.01.
- Utility program autopp added for performing repetitive tasks such as converting files from one format to another or globally replacing a string in a set of files. Note that not all are currently operational. This script should still be considered as under development.

- Fixed bug in PO2R that resulted in ctfmode being hard coded to 1.
- Defocus values in particle parameter files only tested if CTF corrections are used.
- Header always written to auto3dem summary file if file does not already exist, regardless of value of restart flag.
- Minor bug fix to PO2R so that parameter nangle specifying number of steps to be taken along each orientation angle can be set to zero. Doing this allows PO2R to refine particle image origins while leaving orientations unchanged.
- Modify tiff2pif conversion program to optionally accept third command line argument that specifies binning factor. Formerly, users were prompted to enter value.
- Add mrc2pif, pif2mrc, and pif2ccp4 conversion programs. Modify so that programs can either accept file names on the command line or prompt user to enter names.
- Provide options for building applications using static linking.
- Executables now reside in bin/ directory rather than BIN/. To ensure back compatibility, BIN is now a symbolic link that points to bin/.
- Improved instructions provided on using “ppft verbose -1” option in auto3dem input file that is automatically generated by setup\_rmc.
- Move perl modules (.pm files) into modules/ directory and modify perl scripts to look for modules in new location. Replace “use lib do” construct with “use FindBin”.
- Add documentation files that are used by programs robem, oned, ctfDisp, and fixpif.
- Minor cleanup of argument list for subroutine global\_cc in program PPFT. Also affects calling routine global.

## auto3dem v3.01

- Minor modifications to make files to ensure that applications build properly when using Portland Group compilers. The -Mnomain option is now used only with those applications that require access to Fortran libraries, but have main program declared in C source. Remove getarg subroutine from libCommpk to avoid clashes with Portland Group library function. Declare iargc to be of type integer in routine parse\_cmd\_line.
- Add mrc2pif and pif2mrc conversion routines to convert directory.
- Modify calls to XtVaSetValues() from function defocusArrowActivate() in ctfDisp program to avoid segmentation faults on 64-bit machines.
- Cleanup preprocessor directives in robem source.

## auto3dem v3.0



This release gathers for the first time all of the image reconstruction, image preprocessing, and utility codes into a single archive. New additions include robem, ctfDisp, oned, fixpif, tiff2pif, emmap3dt, emmapzoom, emsf, em3dbt, emprj, diffit, normit, zerodens, and a number of conversion routines. The make\_all script now takes two command line arguments

```
make_all (parallel | serial) (all | gui | nogui)
```

The first argument behaves exactly as before, while the second argument specifies which set of applications should be built. In most cases, the second argument should be 'all', but the 'nogui' option may be chosen when doing a build on a machine that does not have the Motif library installed. The 'gui' option is generally not needed, but is provided for added flexibility.

If you are already a user of robem, please be aware that this release contains a major overhaul of the software. Modifications include:

- All Fortran code has been brought up to the Fortran90 standard
- Much redundant, obsolete, or unused code has been deleted
- Old Kernighan and Richie (K&R) style function prototypes have been replaced with prototypes that conform to the ANSI C standard
- Directory structure, make files, and include files have been cleaned up and reorganized
- Minor bugs, mostly related to type mismatches in C function calls, have been implemented
- Routines used by multiple applications have been moved into libraries

In order to build many of the newly added programs, you will need to install the Motif library. We do not currently have a feature in place for automatically determining whether the code is being compiled on a 32-bit or 64-bit architecture. The make.inc.common file is hardcoded for 64-bit hardware and the MOTIFLIB macro will need to be manually edited if you are running on a 32-bit machine.

Running the new programs requires that the EMDIR environment variable be set. The easiest way to do this is to set the EMDIR variable in the .cshrc file and then use it to append the path variable.

```
setenv EMDIR /path_to_programs/BIN
set path = ($EMDIR $path)
```

All make files have been overhauled and a number of subroutines that are common to both the program PPFT (parallel PFTsearch) have been moved into libCommpk. In addition, several of the routines in the PPFT directory have been renamed so as to avoid conflicts with other routines of the same name in libEMF.

fft_map_fill	<i>renamed</i>	fft_map_fill_ppft
map_fft_fill	<i>renamed</i>	map_fft_fill_ppft
map_prj	<i>renamed</i>	map_prj_ppft
map_sym_cavg	<i>renamed</i>	map_sym_cavg_ppft

The following directories have been created to accommodate the newly added code: convert, conway\_tif2pif, ctfDisp, em\_tools, fixpif, lib3DAll, libEMF, oned, pftprj, robem.

## auto3dem v2.04

- Overhaul make files and define library macros in include files. Rename libraries to use more standard naming conventions and rename directories as follows to be more consistent with other Purdue software

PIFlib	<i>renamed</i>	libR
Commpk	<i>renamed</i>	libCommpk
Compap	<i>renamed</i>	libCompar
DIElib	<i>renamed</i>	libDIERCKX
Vfftpk	<i>renamed</i>	libVfftpk

- Add new script flesh\_out.pl to generate complete particle parameter files from files that contain only header lines.
- Fix minor bug in select.pm related to filtering particles on theta and phi
- Get rid of old code delimited by #define OLDWAY preprocessor directives in libPIF.c
- Generate new libPIF.h header file from libPIF.c

## auto3dem v2.03

### *Important notes for building auto3dem*

Auto3dem and the parallel codes that it calls can now be built and run in either serial or parallel mode. An implementation of the MPI library (e.g. mpich) is no longer needed when running auto3dem on a single processor. The make\_all script used to build the executables now takes a single command line argument, with allowed values of 'parallel' or 'serial'. Symbolic links are set by the script as follows

```
% make_all parallel
% ls -g make.inc BIN/mode.pm
lrwxrwxrwx 1 csd357 14 Jun 13 15:14 BIN/mode.pm -> mode_parallel.pm
lrwxrwxrwx 1 csd357 15 Jun 13 15:14 make.inc -> make.inc.parallel
```

```
% make_all serial
% ls -g make.inc BIN/mode.pm
lrwxrwxrwx 1 csd357 14 Jun 13 15:14 BIN/mode.pm -> mode_serial.pm
lrwxrwxrwx 1 csd357 15 Jun 13 15:14 make.inc -> make.inc.serial
```

When building auto3dem for parallel operation, the serial Fortran 90 and ANSI C compilers are called by the corresponding mpiF90 and mpicc scripts. For builds in serial mode, the FC and CC macros in the make.inc.serial file must be manually edited if you are not using the gfortran and gcc compilers. The make\_all script determines whether or not a build of auto3dem already exists. If the mode (serial or parallel) of the current build differs from that of the previous build, all .F files are 'touched' to ensure that they are run through the preprocessor and recompiled.

For auto3dem runs in serial mode, the number of CPUs no longer needs to be specified and any value set using the -ncpu flag is ignored.

Detailed listing of code changes:

- auto3dem.pl and run\_mpi\_prog.pm modified to use mode.pm module. When running in serial mode, auto3dem.pl no longer requires (and quietly ignores) specification of number of nodes. validate.pm module modified to accept mode argument. Version string now specifies whether serial or parallel build is being used.
- make\_all script modified to accept mode (serial or parallel) on command line and set symbolic links to appropriate make include files and Perl modules before initiating build. Make include files simplified to just make.inc.parallel and make.inc.serial.
- Argument intents added to P3DR/exchange\_2\_slab.f
- File extension for P3DR/realtocomplx.F, matrixentries\_slab.F changed to .f
- Makefiles for POR, PCTFR, PSF, PPFT, and PCUT modified to include generic rule for creating objects from Fortran source with .F extension.
- Add preprocessor directives to following files to isolate parallel code and allow optional builds in either serial or parallel mode. Add serial code where needed to replace functionality carried out by MPI routines (e.g. initialization of process number and number of nodes, data reduction). All files with .f extension renamed to use .F extension so that they will be recognized by preprocessor.
  - Compar/bcast\_parameters.f
  - Compar/error\_stop.f
  - Compar/exch\_3d\_1.f
  - Compar/gather3d.f
  - Compar/output\_density.f

- Compar/read\_files.f
  - Compar/read\_map.f
  - P3DR/P3dr.F
  - P3DR/cmpt\_intrps.F
  - P3DR/exch\_intp.F
  - P3DR/exchange\_2\_slab.f
  - PCTFR/Pctfr.f
  - PCTFR/cmpt\_ctf.f
  - PCUT/Pcut.f
  - PCUT/cut\_map.f
  - POR/Por.f
  - POR/cmpt\_ort.f
  - PPFT/global.f
  - PPFT/global\_cc.f
  - PPFT/pftsearch.f
  - PSF/Psf.f
  - PSF/comp\_sfactor.f
- Cleanup and modernization of routines in PPFT
    - Makefile – remove map\_peak\_rest.o from object list
    - map\_peak\_restr.f – remove; functionality now in pftcc\_peak
    - list\_ccs.f - cleanup and intent specification
    - pftcc\_peak.f - major cleanup and reorganization for clarity; move call to rearrange\_fft\_style from here to get\_xy; replace call to map\_peak\_rest with inlined functionality.
    - get\_xy.f - add call to rearrange\_fft\_style after call to ccf\_fft
    - calc\_pfts\_g - add FILT\_FAC to argument list; declare pfft to be complex rather than real and make necessary code changes
    - tsbend.f - general cleanup and intent specification
    - rearrange\_fft\_style.f - rewritten so that rearrangement of array done in place. Single array now passed with intent INOUT and same dimensions as array in calling routine.
    - pftsearch.f - get rid of variables NPHI and NTHE
    - global.f - get rid of variables NPHI and NTHE; specify argument intents
    - key\_info.f - get rid of variables NPHI and NTHE, specify argument intents and completely overhaul routine to improve clarity and readability
    - calc\_mod\_tps.f - rename (NPH, NTH) to (NPHI, NTHE) and declare as local variables rather than subroutine arguments.
    - get\_nview\_mod.f - declare NTHE as local variable rather than subroutine argument.
    - pmap\_fft.f - declare array A to be complex of size NROT rather than real with size 2\*NROT and make corresponding code modifications.
    - fft\_2d.f - specify argument intents and cleanup comments.

- `fft_2d_back.f` - specify argument intents, cleanup comments, and use Fortran 90 array syntax.
- `global_cc.f` - declare array PPFT to be complex rather than real; delete `flip`, `index`, `mode`, and `sq1` from argument list in call to `get_tpo_g`; move hand flip calculations from here into `get_tpo_g`.
- `get_tpo_g.f` - specify intent for all subroutine arguments; delete `flip`, `index`, `mode`, and `sq1` from argument list; move hand flip calculations here from `global_cc`; declare `ppft` to be complex and make necessary code modifications
- `ccf_fft` - Complete overhaul for improved clarity, use of complex arrays, cleanup of logic, commenting. In particular, switching from real to complex arrays vastly simplifies index calculations.
- `fill_params.f` - specify argument intents, extensive overhaul of logic for improved clarity
- `write_params.f` - overhaul and cleanup for improved clarity; specify argument intents.

## auto3dem v2.02

Summary of changes that result in slight numerical differences

- Faster approximate algorithm for determining  $\omega$  in program PPFT. Enabled by default and controlled by 'ppft quick\_omega' parameter.
- Faster approximate algorithm in POR for local search of orientation space. Enabled by default and controlled by 'por quick\_search' parameter.
- Apodized map produced by P3DR. Width of border region controlled by 'p3dr apo\_border' variable, default value equal to 12 pixels.
- Automatic refinement of CTF parameters now turned off by default. Controlled by 'auto refine\_ctf' parameter.
- Bug fix in auto3dem so that inner and outer radii of capsid estimated correctly when binning used in PPFT

Detailed description of changes

- Major performance enhancements made to PPFT routine (`get_phiomega`) responsible for determining the orientation angle  $\omega$  and the sign of  $\phi$ . For smaller problems, this routine accounts for a very tiny fraction of total run time, but owing to the scaling behavior of the algorithm (NROT3, where NROT depends on map size and binning factor), `get_phiomega` can dominate the run time for larger problems.

The initial search for  $\omega$  is now done over a coarser grid of values, with the step size dependent on NROT. This is followed by a local search using a finer grid in the vicinity of the top scoring values of  $\omega$  obtained during the initial

search. After several iterations of global search mode, only a small fraction of the particle images have orientations that differ from those obtained using the original algorithm and the actual differences in the orientations are minimal.

The “quick omega” feature is enabled by default in auto3dem. Adding the following line to auto3dem input files disables this option

```
ppft quick_omega 0 # auto3dem input file
```

This new capability required the creation of a new subroutine `get_phiomega_quick` and modifications to the following files in directory PPFT: `global.f`, `global_cc.f`, `get_tpo_g.f`, `pftsearch.f`, `ppft_info.f`, and `key_info.f`. Also required changes to `include/infohead.inc`, `init_params.pm`, and `make_program_input.pm`

- Capabilities added to P3DR to produce an apodized map, with a Gaussian falloff applied to the density in the region  $\text{boxrad-border} \leq r \leq \text{boxrad}$ . Border has units of pixels and must be an integer. Using a border width of zero recovers original P3DR behavior. Default value in auto3dem set to 12, but can be overridden by adding the following line to the auto3dem input files

```
p3dr apo_border n # auto3dem input file
```

This new capability required changes to the following files: `P3DR/P3dr.F`, `P3DR/density_clear.f`, `Commpk/info.f`, `Compar/bcast_parameters.f`, `info.inc`, `init_params.pm`, and `make_program_input.pm`

- Provide “quick search” capabilities in POR so that a restricted local search of orientation space (one  $\Delta\theta$  step along each direction for each orientation angle) is performed first and only those particle images that find a better orientation in this restricted region are subjected to a more extensive local search. Using the “quick search” feature typically results in a very small fraction of the particle images settling into orientations that are different from those obtained using the more extensive local search, but has been shown to decrease run times by up to a factor of five.

The “quick search” feature is enabled by default in auto3dem. Adding the following line to auto3dem input files disables this option

```
po2r quick_search 0 # auto3dem input file
```

This new capability required changes to the following files: `POR/Por.f`, `POR/cmpt_ort.f`, `Commpk/info.f`, `Compar/bcast_parameters.f`, `info.inc`,

init\_params.pm, and make\_program\_input.pm

- Refinement of CTF parameters using program PCTFR is now disabled by default. To override, add the following line to auto3dem input file

```
auto refine_ctf 1
```

- setup\_rmc.pl prints comments to the auto3dem input file describing how to set parameters to perform faster reconstructions and reach higher resolutions. The following commented out parameters are now listed

```
ppft bin_factor
ppft verbose
ppft annulus_low
ppft annulus_high
pcut in_rad
pcut out_rad
auto freeze_annulus
auto bin_reduce
```

- Bug fix made to Perl module get\_ann\_lo\_hi.pm so that inner and outer radii of capsid are estimated correctly when binning is used in program PPFT.
- POR modified to keep statistics on both the level of individual micrographs and the entire run for the number of particle images that fall into each of the following categories

```
Stable orientation / stable origin
New orientation / stable origin
Flipped hand / stable origin
Stable orientation / new origin
New orientation / new origin
Flipped hand/ new origin
```

The “stable” designation means that the orientation or origin for a particle is unchanged after running POR. “Flipped hand” means that the orientation is the same except for a change in the handedness of the particle. This information is summarized in the POR output file.

Extra fields have been added to the new particle parameter files generated by POR so that the status of the orientation (Stable, New, Flipped) and origin (Stable, New) is listed for each particle.

- Add new Perl module to auto3dem to perform tests that determine whether or not particle parameter files are properly formed. Affects files auto3dem.pl

and sanity\_checks.pm

- Allow for specification of P3DR parameters magfactor (default value 1.0) and map\_dim (default value 0) in auto3dem input file. Affects files init\_params.pm and make\_program\_input.pm
- Modify PCTFR/cmpt\_ctf.f to handle arbitrary formats in the particle records of the particle orientation files. Records are now read in as strings, rather than parsed into individual fields, and written out to the new particle file following the updated CTF parameters. This allows us to include additional fields that may be written by programs POR or PPFT without having to make modifications elsewhere in the code.
- Command line syntax for auto3dem and setup\_rmc more flexible. Key value pairs no longer need to be separated by equal signs. For example, both of the following are now valid

```
%auto3dem -ncpu=16 -input=input_file
```

```
%auto3dem -ncpu 16 -input input_file
```

- Minor performance enhancements made to  $(\theta, |\varphi|)$  search routine get\_thephi in program PPFT.
- Make following modifications to PPFT input routine ppft\_info:
  - Equal sign in key-value pairs is now optional
  - Strings containing '/' properly handled
  - lg\_out initialized to zero
  - Start of data file list can be specified using both 'end\_of\_keys' (first three characters significant) or 'inputparameterfiles' (first nine characters significant).
- Allow blank lines and comments (using '#') in the data file section of input for programs PPFT, POR, P3DR, and PCTFR. Required changes to PPFT/pfile\_info\_pft.f and Commpk/pfile\_info.f.
- The remaining modifications listed below for v2.02 all involve internal code changes that do not affect results, use of codes, or specification of input parameters.
- Cleanup timing functionality throughout code
  - Add new function elapsed\_time.f to library Commpk that calculates the time elapsed between pairs of system\_clock readings; get rid of old timer function timer1.c, and update Makefile.
  - Modify mem\_time.inc so that all timing variables reside in a common block and get rid of test\_memory, test\_time, and wtime declarations.



- Replace calls to `mpi_wtime` with Fortran 90 `system_clock` intrinsic in the following files: `Ctffit/ctffit.f`, `P3DR/P3dr.F`, `P3DR/cmpt_intrps.F`, `PCTFR/Pctfr.f`, `PCTFR/cmpt_ctf.f`, `PCUT/Pcut.f`, `POR/Por.f`, `POR/cmpt_ort.f`, `PPFT/global.f`, `PPFT/global_cc.f`, `PSF/Psf.f`
- Get rid of old commented out code in following files: `P3DR/matrixentries_slab.F`, `PCTFR/Pctfr.f`, `POR/Por.f`, `POR/cmpt_ort.f`.
- Comment out or remove unused variables and labels in following files: `PCTFR/cmpt_ctf.f`, `PPFT/list_ccs.f`, `PPFT/pfile_info_pft.f`, `PPFT/ppft_info.f`, `Ctffit/ctffit.f`, `Ctffit/rotate.f`

## auto3dem v2.01

- Modify programs `PSF`, `PCUT`, `P3DR`, `PCTFR`, `PPFT` and `POR` to allow specification of input file both as command line argument and through redirection of standard input. For example, both of the following are valid ways to execute `P3DR`

```
% mpirun -np 8 P3DR p3dr_input > p3dr_output
% mpirun -np 8 P3DR < p3dr_input > p3dr_output
```

The first syntax is preferred and is used internally by `auto3dem` when launching the parallel jobs. The reason for making this change is that many MPI implementations cannot handle redirection from standard input if the number of characters is greater than 4096. The second syntax is retained for continuity purposes only and is no longer recommended.

- Add full path to nodefile in `setup_rmc` and `auto3dem`.
- Add capabilities to change search mode bin factor in `auto3dem` using tests analogous to those used in making decision to switch from search mode to refine mode. Controlled through 'auto bin\_reduce' parameter. Bin factor is now listed in summary file.
- Fix logic error in `auto3dem` that resulted in mode (search → refine) being updated after the restart file was written. This bug did not affect results of image reconstruction, but could result in an extra iteration of search mode being carried out if the calculations died during the first iteration following the automated transition from search to refine mode.
- Added new script `config_test.pl` to main code directory to perform simple tests on the computing environment. Display Perl version is displayed and determine whether or not all required and optional Perl modules are present. Confirm that `mpirun`, `mpicc`, `mpif90`, and scalar C and Fortran compilers called by `mpif90` and `mpicc`, respectively, are in path.

The remaining modifications listed below for v2.01 all involve internal code changes that do not affect results, use of codes, or specification of input parameters.

- Eliminated temporary array PRJSL in file PPFT/global.f
- Cleanup and reorganization of PIFlib/libPIF.c

## auto3dem v2.0

**Important** - Version 2.0 contains a new command line interface for running `setup_rmc` and `auto3dem`. The old usage syntax is no longer valid. Any scripts that call `setup_rmc` or `auto3dem` must be modified to use the new syntax.

- Completely overhauled command line interface for both `auto3dem` and `setup_rmc` so that input is done using key-value pairs. For example

```
setup_rmc -ncpu=4 -seed=123 -list=listfile
setup_rmc -usedefaults
```

```
auto3dem -ncpu=12 -input=input_file -nodefile=mynodelist
```

`setup_rmc` can be run purely with default values using the `-usedefaults` flag. For `auto3dem`, values must be provided for the number of CPUs and the name of the input file. The nodefile no longer needs to be specified for either `setup_rmc` or `auto3dem`. This is to be contrasted with previous versions where the word “none” had to be explicitly used for those cases where the parallel computing environment did not require a node list. For batch systems running the PBS scheduler, the PBS node file is automatically obtained from the `$PBS_NODEFILE` environment variable via Perl’s `%ENV` hash.

Both programs list usage information if executed either without any arguments or with the `-help` flag. All flags are case insensitive and whitespace before or after the “=” in the key value pairs is tolerated. In addition, `setup_rmc` prompts the user to continue run using all default values after printing usage information.

- Major performance improvements in P3DR for icosahedral symmetry. Twofold and threefold symmetry operations are applied to 3D DFT of model, thereby reducing the number of interpolations required for each image from 60 to 5. Speedup is problem and system dependent since it depends on the number of particle images and the relative time required for different operations (FFT, interpolation, etc.), but runs times have been measured to be reduced by 8x for test cases involving 1000 particle images.

- Added capability to handle both relative and absolute path names to boxed image files (as specified in particle parameter files) to setup\_rmc.
- CTF corrections can be turned off globally using -noctf option on the setup\_rmc command line and “auto noctf 0” in the auto3dem input file. This has the same effect as manually setting the CTF mode for programs P3DR, PO2R, PCTFR, and PPFT to zero. Using the -noctf option in setup\_rmc automatically propagates setting of the noctf flag in the generated auto3dem files.
- The auto3dem input file now recognizes the keys ctf\_mode and ctfmode for all programs that perform CTF corrections. This change was made to remedy the confusing situation where PPFT used ctf\_mode while P3DR, PO2R, and PCTFR used ctfmode.
- Intermediate files generated during random model method calculations are all moved to a new directory named RMC\_temp. The best starting model is renamed rmc.pif and moved into the directory containing the particle parameter files. Files and directories corresponding to random orientations that are not used are automatically deleted using a new utility script, remove\_useless.pl, which is called from RMC\_run. A short script named RMC\_cleanup is generated by setup\_rmc for the purpose of recursively removing contents of RMC\_temp.
- setup\_rmc automatically generates a basic auto3dem parameter file to continue reconstruction after random model method calculations have been completed.
- Sensible defaults now used for obtaining list of particle parameters files considered by setup\_rmc. If -list key is not used, setup\_rmc first looks for a file named 'list'. If this file does not exist or is not readable, then the data directory is queried for files of the form \*000, \*001, etc.

When a list of parameter files is read from a file, now have added flexibility in the specification of the files. Wildcards, comments, blank lines, and whitespace are now allowed. Can also use a single quoted expression in place of the file. For example, -list='\*001' will be expanded within setup\_rmc to a list of all files ending with 001 in the specified directory.

- Modified auto3dem.pl so that GD::Graph module is loaded at run time using “require” rather than “use”. This makes it possible to trap exceptions and bypass the FSC graph generation if the module cannot be found. The advantage of this approach is that the auto3dem.pl source does not have to

be manually edited to handle Perl installations that are missing GD::Graph.

- PCTFR is now called by auto3dem only if 'auto mode' equals refine and both 'pctfr ctfmode' and 'auto refine\_ctf' are true (non-zero).
- Lower memory algorithm and single precision interpolation option for P3DR have been removed. This both simplifies the source code and makes it easier to implement planned performance enhancements. The default is to use the more memory intensive (and less communications intensive) algorithm and double precision interpolation options, but these could previously be overridden using the -DALG3 and -DSPREC compiler flags.

An older version of P3DR has been retained in the P3DR\_old directory, but does not contain the performance enhancements described above.

- Semicolons now allowed, in addition to commas and whitespace, as delimiters in list of email addresses in auto3dem input file. For example,  
  
auto recipient address1 address2, address3; address4
- MPI communications in PPFT/global.f modified to avoid exceeding MPI buffer limitations. Allows PPFT to be run using a finer spacing between projections of the model and/or larger models.
- MPI communications in P3DR/exch\_intp.F modified to avoid exceeding MPI buffer limitations. Allows P3DR to be run for larger image sizes.
- PSF/Psf.f modified to give warning message rather than terminating execution if the pixel sizes in the even and odd map headers are different.
- PCTFR/Pctfr.f modified to use pixel size as specified in the particle parameter files rather than the map files. Warning messages are provided if the pixel size of the map is either negative or larger than the particle parameter file pixel size.

The remaining modifications listed below for v2.0 all involve internal code changes that do not affect results, use of codes, or specification of input parameters.

- Variable r4\_max set to huge(r4\_max) in PCTFR/cmpt\_ctf.f and POR/cmpt\_ort.f
- General cleanup and improved commenting of Compar/bcast\_parameters.f and Compar/read\_files.f
- Remove nodefile from validate\_commandline\_args() argument list in validate.pm and from call to function in auto3dem.pl. Allow nodefile to be undefined in auto3dem, writer\_header, and run\_mpi\_prog.

## auto3dem v1.12

- Default values are now set for the “auto outfile” and “auto rundir” parameters. Unless otherwise specified, outfile is set to the name of the working directory where auto3dem is launched and rundir is set to be the directory dat. This involved changes to the file init\_params.pm.
- Modified Fortran input routines pfile\_info and pfile\_info\_pft to handle more flexibility in the specification of the image file names in the particle parameter files. The image name is first tested exactly as entered in the parameter file. If the program is unable to open it, the path name to the file is stripped off and an attempt is made to open the file in the current working directory. Similar changes made in the Perl module get\_info.pm.
- Rename the auto3dem parameters fsc\_locut and fsc\_hicut to fsc\_lothresh and fsc\_hithresh, respectively. Scripts will still accept the old parameter names, but will be converted internally to new names. These changes were made so that the names of the variables would more accurately reflect their usage as thresholds in the FSC data. Similarly named variables in the Perl code were renamed to use the \_lothresh and \_hithresh suffixes. Also reordered return argument list from PSF parsing routine. These changes involved modifications to the files init\_params.pm, auto3dem.pl, psf\_parse.pm, and updatre\_res.pm.
- Minor bug fixes to the files Commpk/ctf\_para.f, POR/Por.f, and P3DR/P3dr.F to properly handle the case where no CTF correction is applied. PCTFR/Pctfr.f modified so that it exits early if ctfmode is set so that no correction is applied. Commpk/info.f fixed to properly handle num\_den\_pix set equal to zero in the input files.

The remaining modifications listed below for v1.12 all involve internal code changes that do not affect results, use of codes, or specification of input parameters.

- Intent of interpolants argument in P3DR/cmpt\_interps.F changed to INOUT.
- In the following routines, replace the variable stdoutoutput with \* in write statements: Commpk/intlz\_arrays.f, Commpk/symmcode.f, P3DR/P3dr.F, P3DR/exchange\_2\_slab.f, P3DR/fftsynth\_1\_m\_slab.f, P3DR/fftsynth\_1a.f, PCTFR/Pctfr.f, PCUT/Pcut.f, and POR/Por.f
- Completely overhaul the include file include/allprog.inc. Get rid of stdinut, stdoutoutput, max\_input, pi, twopi, deg\_to\_rad, and rad\_to\_deg; use Fortran90 parameter statements to set constants; improve comments; move filename\_len from allprog.inc to info.inc. Get rid of variable skipone since it is used only in readorient.f and declared with SAVE attribute to retain value

between calls.

- Cleanup and overhaul include file include/info.inc. The variable filename\_len is now declared here so that the use of info.inc does not rely on allprogs.inc.
- Get rid of skipone variable in Commpk/info.f
- Replace filename\_len with hardcoded value 257 in files Compar/read\_map.f and Ctffit/ctffit.f
- Get rid of oddeven parameter in POR/cmpt\_ort.f and PCTFR/cmpt\_ctf.f; in calls to readorient, replace with zero.
- In routine Commpk/intlz\_params.f get rid of initialization of pi, twopi, rad\_to\_deg, and deg\_to\_rad
- Overhaul logic in PPFT routine involved with the VAROPT feature. This is an experimental feature that is currently not recommended, but code modifications may have fixed a long time bug. Changes involved files global\_cc.f, pftcc\_fill\_g.f, get\_tpo\_g.f, avg\_pftimg.f, and get\_thephi.f. Also declare variable pi locally in pftsearch.f

## auto3dem v1.11

- Email notification feature modified so that messages can be sent to multiple email addresses and that an arbitrary number of files can be attached. Individual addresses may be separated by whitespace and/or commas and each individual address is tested to be well formed. Note that additional fields in auto3dem parameters file records are still ignored except for the “auto recipient” record. Modifications involved files auto3dem.pl, sendmail.pm, and init\_params.pm
- Capabilities added to generate and email graphs of FSC curves using GD library. Since not all sites have the required GD and GD::Graph Perl modules, the code in auto3dem.pl marked with the #GDSFC to construct the graphs must be manually uncommented.
- Input files info.f and pfile\_info.f in library Commpk completely overhauled. Can now handle P3DR and PO2R input files that contain blank lines, tabs and leading whitespace.
- OPEN statements in key\_info.f and write\_params.f modified so that PPFT no longer dies if it tries to overwrite existing files.

- Added subroutine indexx (borrowed from Numerical Recipes in Fortran90, with very minor modifications) to the Commpk library. Will be used in future releases to sort records in particle parameter files by their ID.
- PCTFR/cmpt\_ctf.f and POR/cmpt\_ort.f remove all whitespace from particle image file names before writing particle parameter files.
- POR/cmpt\_ort.f now writes zeroes for the last two fields in the particle parameter files. By making this change, the files generated by PPFT and POR will have the same format.
- Commpk/readorient.f modified so that it can deal more reliably with particle parameter files that are missing the last two scores. These types of files are generated by OOR and older versions of PO2R.
- File Commpk/read\_1\_pif.f cleaned up and commented. Include statements removed and argument list expanded. Corresponding changes made to argument list in Por.f, P3dr.f, and Pctfr.f.

Minor changes made to auto3dem so that:

- Temporary files stdin\_temp and message.txt are deleted
- Log and summary files are properly handled in the event that a new run is launched but the log and/or summary files already exist
- FSC data and graphs labeled using same prefix as map, summary, log, and restart files.

## auto3dem v1.10

- setup\_rmc no longer specifies the number of models on the command line. New syntax is

```
setup_rmc dir list ncpu nodefile [boxrad]
```

- Streamlined random model calculations – auto3dem.pl, init\_params.pm, write\_params.pm, setup\_rmc.pl, and findbest.pl were modified so that calculations on a particular random model are terminated once the resolution improves to the point where the FSC curve never drops below 0.5. Also, additional random models are not calculated once one of the random models had successfully converged. In the best cases, this results in a 20x reduction in the run time relative to calculating ten full iterations for ten different random models.

- Maps generated by auto3dem have a better naming convention, using the “auto outfile” parameter to specify the file prefix and “.pif” to specify the suffix. For example, the 5th map generated where outfile equals reovirus will now be named reovirus\_iter\_5.pif rather than map\_iter\_5. Script findbest.pl modified to handle new naming conventions.
- Improved naming of variables in several of the Perl scripts and modules.
- Always output memory usage and timings in P3DR, PCTFR, PCUT, and POR. These had been under control of the hard-coded logical variables test\_time and test\_memory.
- Get rid of calls to bcast\_parameters in Psf.f and Pcut.f. This makes the code easier to follow since bcast\_parameters broadcast parameters that were listed in a common block and allows for more sensible naming of variables.
- Get rid of “include ../info.inc” in PCUT/cut\_map.f and expand argument list. Get rid of “include ../mem\_time.inc” in Pcut.f and declare necessary variables rather than relying on common blocks.
- Completely overhaul logic in PSF to make the sequence of operations easier to follow.

## auto3dem v1.09

- Email notification – a new module sendmail.pm was added to AUTO3DEM that provides capabilities for sending email notifications with optional attachments. This feature requires that the standard Linux/UNIX mutt email tool be installed on your system. The sendmail.pm module tests for mutt using the UNIX ‘which’ command and does a cursory check on the email address to make sure that it meets the minimal requirements for being well formed. auto3dem.pl has been modified to send email notifications, with a text output of the FSC calculations, at the end of each iteration. To enable this feature, add the following line to the AUTO3DEM parameter file, where email is a valid email address.

auto recipient email

- Random model computations have been simplified. Modifications have been made to both setup\_rmc.pl and findbest.pl so that best map is identified and copied into the directory provided as argument to setup\_rmc. Argument list for setup\_rmc has been expanded so that the commands file does not need to be hand edited and so that all temporary files and directories associated with random model computations have the ‘RMC’ prefix. This last change makes



the cleanup of intermediate files more straightforward. The updated syntax for running setup\_rmc is shown below. Note the new ncpu and nodefile arguments.

```
setup_rmc dir list nmodels ncpu nodefile [boxrad]
```

- CTF refinement – the program PCTFR (Parallel CTF Refinement) has been integrated into AUTO3DEM. This involved changes to auto3dem.pl, init\_program.pm, make\_program.pm, and write\_param.pm. The subroutine cmpt\_ctf called by PCTFR has been overhauled and modified so that it can handle both stigmatic and astigmatic images. PCTFR does not currently handle the transition from a stigmatic to astigmatic images unless a reasonably accurate estimate of the orientation of the astigmatism is provided. The main program Pctfr.f has been cleaned up and modernized.

CTF refinement is controlled by the new parameter 'auto refine\_ctf' and is enabled by default. Note that CTF refinement is only done when running in refine mode since at lower resolutions the spatial frequencies may not span enough nodes in the CTF function to make a reasonable estimate of the defocus.

- Multiplication of model projections by CTF in PO2R – In previous versions of PO2R, the projections of the model were divided by the CTF function before comparing them to the images. Modifications were made to files Por.f and cmpt\_ort.f so that projections of the model are multiplied by CTF. Either approach is valid for comparing the model to the images, but multiplication by the CTF avoids the approximations that are required to properly handle the inverse CTF in the vicinity of the nodes.
- Tests added to AUTO3DEM (module update\_res.pm) to ensure that the maximum resolution used in P3DR is not less than twice the pixel size (Nyquist limit).
- Default values now set for PCUT in\_rad and out\_rad; PPFT annulus\_low and annulus\_high. It should be noted though that it is still to the user's advantage to manually enter values for these parameters once estimates become available from lower resolution maps.
- PPFT modified so that existing output files are overwritten rather than having the program terminate. In addition, PPFT can now handle leading whitespace before the names of parameter files and trailing empty lines in the input file.
- General code improvements – removed unnecessary include statements in fft\_2dfft; generalized maptempfac to handle arbitrarily shaped arrays;

modified readorient to return score and correlation coefficient data from files generated by PO2R and PPFT, respectively; modified cmpt\_intps, cmpt\_ctf, and cmpt\_ort to be compatible with new readorient routine;

- Temperature factor – ctf\_para now applies temperature factor in the same manner (positive sign in exponent) for all values of filter. Simplifies logic of calling ctf\_para with regards to sign convention. All code (except possibly PPFT) should handle temperature factor calculations correctly.
- Random model pixel binning – changed threshold for using binned data (bin\_factor=2) in random model calculation from pixels of size 5Å to 4Å.
- Hollow map bug fix – Repaired bug that affects hollow map calculations when running in search mode. Since hollowed maps are an advanced feature and are normally only used in refine mode, this bug probably had no impact on any reconstructions performed to date.
- PO2R res\_min bug fix – Repaired bug so that PO2R minimum resolution is properly calculated. Bug would only have been encountered if user was trying to manually override default value for res\_min based on calculation involving pixel size.
- PO2R handedness test bug fix – Repaired bug that had been introduced in version 1.08 regarding test for proper hand.

## auto3dem v1.08

Modifications dealt primarily with the elimination of redundant functionality. In particular, the use of equivalent\_view to generate symmetry-related orientations and crowther\_to\_matrix to determine rotation matrices corresponding to the orientations defined by (theta, phi, omega). These changes lead to very small numerical differences. A number of non-algorithmic modifications were also implemented that improve code readability.

- Commpk – split the file to\_asym\_unit.f into separate files each containing a single subroutine or function; got rid of common block and used SAVE attribute to retain the constants in function good and subroutine genrot that were previously calculated in geometry\_init. Removed subroutines setrotmat and getorient since these are no longer needed. Moved the routines density\_clear and maptempfac from library Compar to Commpk since they do not contain any parallel code.
- Compar – Moved routines density\_clear and maptempfac to library Commpk.

- P3DR – Overhauled calculation of symmetry-related orientations for icosahedral symmetry in subroutine cmpt\_intrps; replaced a fairly large block of confusing code with calls to routine equivalent\_view. Replaced call to setrotmat with crowther\_to\_matrix.. Updated calling sequence to eight\_symmetry and dihedral symmetry to take (theta, phi, omega) as arguments rather than a 3-element vector. Got rid of subroutine ico\_vector since it is no longer needed for determination of equivalent orientations.
- POR – Replaced calls to setrotmat with crowther\_to\_matrix in subroutine cmpt\_ort.
- PCTFR - Replaced calls to setrotmat with crowther\_to\_matrix in subroutine cmpt\_ctf.

## auto3dem v1.07

- PPFT – get rid of files wavel.f, ctf\_scale.f, and piraddeg.f; move ctf\_firstpeak\_pft.f to library Commpk, and replace calls to ctf\_scale with calls to ctf\_para in subroutines global\_cc and calc\_pfts\_g. Make minor, non-algorithmic (general cleanup, redefining real arrays as complex arrays, specification of argument intents, etc.) changes to multiple files.
- POR – change one-dimensional arrays to multi-dimensional arrays in Por.f and cmpt\_ort.f. This does not change any results, but does make code easier to follow. Get rid of file imgcompctet.f and replace calls to imgcompctet with calls to imgcomport.
- P3DR - change one-dimensional arrays to multi-dimensional arrays in P3dr.F; minor modifications to matrixentries\_slab.F.
- PCUT - change one-dimensional arrays to multi-dimensional arrays in Pcut.f
- PCTFR - change one-dimensional arrays to multi-dimensional arrays in cmpt\_ctf.f; minor modifications to Pctfr.f
- PSF - change one-dimensional arrays to multi-dimensional arrays in Psf.f
- Compar – general cleanup of exch\_3d\_1.f
- Commpk – split ctf\_para.f file into multiple files ctf\_es.f, ctf\_et.f, ctf\_firstpeak.f, ctf\_func.f, ctf\_para.f, and ctf\_temp.f. Overhaul and cleanup all routines, with an emphasis on making them usable by POR, P3DR, and PPFT. Move PPFT/ctf\_firstpeak\_pft.f into Commpk. Perform general cleanup of

disef\_para.f, focus\_astig.f, imgcomport.f, intlz\_arrays.f, info.f, rpifimag.f, symmcode.f

## auto3dem v1.06

- Remove files ico\_EM4IMR.f and ico\_EM4IMR\_vector and replace with single routine in file ico\_vector.f. New version may lead to very slightly different results since the hard-coded values 0.809017, 0.500000, and 0.309017 have been replaced with the more accurate representations  $\cos(36^\circ)$ ,  $\cos(60^\circ)$ , and  $\cos(72^\circ)$ , respectively.
- Modified the code for filter=1 in ctf\_para so that CTF is zeroed in when the condition  $|CTF| < 0.1$  is met rather than  $|CTF| < \sqrt{ctf\_ff2}$ . The old version could possibly lead to the CTF being set to zero for a very large range of CTF values.
- Removed the include file vax\_minmax.inc since it is no longer needed by any routines.

Minor, non-algorithmic changes made to many files in Commpk, Compar, P3DR, POR, PPFT, PSF, include, and Ctfrit directories. Primarily removal of unused variables, specification of argument intents, improved comments, etc.

## auto3dem v1.05

- Improved capabilities added for exiting auto3dem if one of the image processing programs aborts. Depending on operating system and batch queuing system, the return value from launching an MPI job with system command may not be accessible to the auto3dem script. To ensure that the abort is detected, the output from the MPI program is parsed and auto3dem exits if the string MPI\_Abort (not case sensitive) is found. Involved modifications to error\_stop.f, run\_mpi\_prog.pm, and auto3dem.pl.
- Minor performance improvements made to 3D interpolation routine interpl\_3d.
- General cleanup of cmpt\_intrps.F, fftsynth\_1\_m\_slab.f, exchange\_2\_slab.f, P3dr.F, Pcut.f, ctf\_para.f, Psf.f, Por.f, Pctfr.f, com\_01.inc, intlz\_arrays.f, and intlz\_params.f. Specification of argument intents, removal of unused variables, etc.
- General cleanup of PSF/comp\_sfactor.f. Move calculations contained within subroutine cplex2ap into comp\_sfactor and get rid of unnecessary operations.
- New control parameters added:
  - auto switch\_mode (see note below)

- auto term\_refine (functionality not yet active)
  - auto term\_search (functionality not yet active)
- Auto3dem now monitors resolution and can automatically switch from search to refine mode if the following conditions are met:
  - switch\_mode flag is true
  - Resolution has not improved by at least 0.25Å over previous iteration
  - Mode is currently search
- Iteration information is now tracked so that numbering from one run to the next is maintained. Restart and continuation files specify the last iteration that had been completed. The restart files also keep track of the number of iterations required to finish the original calculation.
- For restart or continuation, new results are appended to the log and summary files.
- Restart files are written at two points in each iteration: after the completion of origin and orientation refinement and after construction of map. The restart files properly set the auto have\_map flag so that map is constructed from particle images if program needs to be restarted after first checkpoint.
- Output files are named consistently using the name of the auto outfile parameter:
  - outfile\_log – detailed output log
  - outfile\_summary – summary information
  - outfile\_restart\_na – restart after first checkpoint, iteration n
  - outfile\_restart\_nb – restart after second checkpoint, iteration n
  - outfile\_continue – continuation file for successfully completed runs
- Summary file now lists the number of particles that had been selected to construct the final map and the number of CPUs (MPI processes) used.
- Format of restart and continuation files has been improved to logically group parameters.
- Default value (0.1 pixels) specified for po2r dcenter, step size for origin refinement.
- auto final\_map no longer used and has been labeled as deprecated.

## auto3dem v1.04

- Auto3dem version now printed in log and summary files
- Symbolic links to executable Perl scripts in BIN directory created. If path information is setup correctly, can use name of script with or without .pl extension.
  - auto3dem *points to* auto3dem.pl
  - setup\_rmc *points to* setup\_rmc.pl
  - handflip *points to* handflip.pl
  - findbest *points to* findbest.pl
- auto3dem and setup\_rmc executed without any arguments provides usage information and version

- Improved format for auto3dem summary data
- Fixed bug in auto3dem.pl where PPFT output files were moved before they were no longer needed.
- Minor changes to output from update\_res.pm
- Default value for lower resolution limit of FSC curve (psf res\_min) set to 60Å
- Default value for PO2R angular step size (po2r dangle) set to one-half of the angular step size for PPFT.
- Default value for PO2R res\_min set to  $2/5 * \text{boxrad} * \text{pixelsize}$
- Default value for PPFT resolution\_low set to  $2/5 * \text{boxrad} * \text{pixelsize}$
- Test provided on number of iterations in input file. Must be defined and be at least zero.
- Test added to make sure that at least one 'data' line is specified and that if wildcards are used they expand to include at least one file.

## auto3dem v1.03

Changes were implemented that make it much easier to run AUTO3DEM. Rather than creating symbolic links to Perl code and binaries, path information is simply added to the computing environment. This required some changes to the Perl code so that binaries and custom modules would always be found.

### auto3dem

- Custom Perl modules reside in the same directory as the executable Perl scripts. Added the following line to all .pl and .pm files so that custom modules would always be found, regardless of where the scripts are executed

```
use lib do { __FILE__ =~ m|^(\.?)*[/]*$; "$1"; };
```

- Modified setup\_rmc.pl so that executable Perl scripts in the command file are not pre-appended with "./". For example

*./auto3dem.pl becomes auto3dem.pl*

- In module init\_params.pm, get rid of full path information for binaries

### make\_all

- Modified so that all binaries are copied into the BIN directory

## auto3dem v1.02

### auto3dem

- Modified auto3dem.pl and init\_params.pm so that map can be constructed without doing resolution estimation. This feature is particularly useful when applying inverse temperature factor to final map.
- Particle selection criteria can now be applied globally across all files or on a per file basis. By default, selection criteria applied globally. Choice controlled by auto\_global\_select.
- Create subdirectory for storing the program input and output files, together with the filtered particle parameter files generated by applying the particle selection criteria.
- Rename the PSF output file corresponding to the best particle selection criterion for a given iteration to FSC\_curve\_{iteration}.
- Handle the case where the FSC curve never drops below fsc\_locut. Particular useful in the early stages of search mode where resolution of map has the most rapid improvements.
- Get rid of dryrun option
- Write summary file containing just the most important information required to monitor the progress of the reconstruction.
- Modify init\_params.pm to ignore 'auto dryrun' and 'auto split\_ppft' and provide message that these are deprecated options.

#### **Program POR**

- Modified Por.f so that program can deal with map files that contain ridiculous values for the pixel size. In this case, the pixel size from the particle parameter file is used instead.

#### **handflip.pl**

- Script for changing the handedness of orientations in particle parameter files.

### **auto3dem v1.01**

#### **Program P3DR**

- Interpolation routine matrixentries\_slab.F improved to yield better performance. Actual speedups will vary by problem and hardware, but runs using test problem show 28% reduction in time spent in 2d interpolations and a 15% reduction in overall run time.

#### **Program P3DR**

- Modified P3dr.F and get\_1st\_ortid.f to properly handle empty particle parameter files.

#### **Program PPFT**

- Added option for verbose=-1, which results in the correlation coefficient cc\_cmp not being calculated. This option should only be used for construction

of starting model using random model method since cc\_cmp is generally used to select particle images.

### **setup\_rmc**

- Script setup\_rmc.pl modified so that default values for PPFT are overridden.

delta\_theta = 1° (default 0.5°)  
verbose = -1 (default 2)

### **auto3dem**

- parameter file now accepts po2r for specifying PO2R input. por still works to allow back compatibility with v1.0, but all internal perl variables renamed to use po2r (*e.g. %por\_ref renamed internally to %po2r\_ref*)