APPENDIX L

WHO VERSION 5

Instructions for running the SPAT using the WHO5 system are given below. WHO5 is a modified version of the WHO4 system developed by Sadler (2012). It was modified to accommodate a larger data set of an arbitrary number of objects, to implement the correction for peak-to-tail difference among objects of different brightness, to increase double the number of differential points by modifying the interpolation scheme, and to fine-tune the interpolation uncertainties.

The operating system used was openSUSE. The scripts were written in the tc shell, and the FORTRAN compiler used was the Portland Group FORTRAN 90 compiler (pgf90). If using another shell or a different compiler, modifications to the scripts will be needed. The plot package Gnuplot was used for creating all plots, except for Fig. 3.9, which was created by a combination of NumPy and Matplotlib Pyplot in Python. G3data [1] and online tool WebPlotDigitizer [2] were used interchangeably for extracting data from published plots, wherever a plot was adapted from literature.

L.1 Main components

Table L.1: These components must be present in the working directory to run the code. This description largely follows (Sadler, 2012) with several additions and modifications for this dissertation.

Component	Function
DATACHECK_NEW1	This code checks if there is LC data in region II for all SNe Ia, except for CSP-I. If an object does not have data in that region, then its name is written in a file named "BadSNeAuto". All other usable object names are written in the "in.dat" file, and their redshifts in the "redshifts25.dat" file. The numbers of the usable SNe are written in the "input.par" file.

[1] `https://directory.fsf.org/wiki/G3data`
[2] `https://apps.automeris.io/wpd/`

Table L.1 - Continued

Component	Function
NEWCHECK	Removes the duplicate or same-day data points from LC data.
ATEST (Sadler, 2012)	"Finds optimal primary parameters based on V-band LC flat-window. Writes output to "in.dat.opt" file".
DATACHECK_NEW2	This code checks if the LC data covers up to +25 days. If not, then the name of the object is written in the "BadSNeAuto2" file. It also rewrites the "in.dat", "redshifts25.dat", and "input.par" files.
NARC (Sadler, 2012)	"If LCs with large gaps are found then this code fills them with template data".
ATEST1_tcsh	"Calls and recalls ATEST using augmented data from NARC if LCs have large gaps around maximum light" (Sadler, 2012). Rewrites "in.dat.opt".
CONSOLID	"Combines "in.dat.opt" files when augmented data is needed (when using NARC)" (Sadler, 2012).
SADIE013 (Sadler, 2012)	"Calculates second-stage primary parameter modifications".
BANDER2v (Sadler, 2012)	"Re-produces "in.dat.opt" with the same s-correction as the V-band from ATEST1, but finds t_0 and m_0 based on quadratic spline fits".
BANDER2u (Sadler, 2012)	Same as BANDER2v, but special for u-band, not used for the analysis in this thesis.
ALPHA11 (Sadler, 2012)	"Conditions the raw LC data with the primary parameters (stretching, shifting, interpolating over uniform timesteps)".
SADIE (Sadler, 2012)	"Determines signal presence of differentials in shifted and stretched LCs".
UDEV (Sadler, 2012)	Computes $(u - V)$ color.
EPSILv (Sadler, 2012)	Guides principal components (PC) finder in V-band.
EPSILu (Sadler, 2012)	Same as EPSILv, but in u-band.
EPCSOLV3v (Sadler, 2012)	Solves overdetermined system of PC coefficients, used for PCr (residuals) option.
EPCSOLV4v (Sadler, 2012)	Same as EPCSOLV3v, but for PCs of differentials (PCd option).
EPCSOLV3u (Sadler, 2012)	Same as EPCSOLV3v, but for u-band.
EPCSOLV4u (Sadler, 2012)	Same as EPCSOLV4v, but for u-band.
INCLUDE	This folder contains all header files.

Table L.1 - Continued

Component	Function
CODES	This folder contains all source codes, except the source codes for projection of the generic parameters to the physical space.
Peter	This folder contains the source codes for the projection of the generic parameters to the physical space, as well as the theoretical relationships to project them ("mrho.data" file). It also contains the names of the "messenger" SNe.
names_CSPI8.dat	Lists the names of the eight messenger SNe Ia from CSPI. This file must be present in the folder named "Peter".
RDAT	This folder contains all raw LCs. The names are in the format "xxxxxx.v.fin" (must be 12 characters), where the "xxxxxx" (first six characters) is the name or abbreviated name of the SN, "v" is for the V-band, and is substituted by "b" and "u" for the B-band and u-band, respectively. "The last three characters must be "fin". The data files must have a three-line header and have three columns of data. These three columns of data, from left to right, are time in days, apparent magnitude, and the photometric uncertainty of the magnitude. The length of the data files can be as needed", (Sadler, 2012).
SDAT - same as Sadler (2012)	"This folder contains temporary data secondary parameter LC signals".
TDAT - same as Sadler (2012)	"This folder contains templates for s-correction normalization and interpolation. The templates that must be in this folder are "ntemp.ti", "btemp.ti6", "utemp.ti6"."
RUNS	Contains tarballs of the runs.
temp.dat (Sadler, 2012)	"File lists the names and Δm_{15} values of the templates in TDAT folder".
names_CSPI.dat	File contains names of the CSP-I SNe Ia in the sample.
in.dat_all	File contains the names of all SNe Ia used in a run. The format is "xxxxxx.y.fin", see above.

Table L.1 - Continued

Component	Function
redshifts25.dat_all	File contains the redshifts of all SNe Ia used in a run. The names are not necessarily in the same order of "in.dat_all". Each line must have the SN-name (in the same format as "in.dat_all", except the last three characters will be "raw" instead of "fin"), followed by a redshift value in the floating point.
WHO5	Modified master script, the new version of the previous WHO4.

L.2 Plotting routines

Descriptions of the plotting routines in the system and instructions for using them are given below. All plotting routines are in the main folder SN-DATA-RUN unless otherwise mentioned.

1.
   - **LCP3v_d1:** Modified from LCP3v script of Sadler (2012). Creates overlay plots of the stretched and shifted LCs - batch 1.
   - **LCP3v_d2:** Modified from LCP3v script of Sadler (2012). Creates overlay plots of the stretched and shifted LCs - batch 2.
   - **LCP3v_d3:** Modified from LCP3v script of Sadler (2012). Creates overlay plots of the stretched and shifted LCs - batch 3.
   - **LCP3v_d4:** Modified from LCP3v script of Sadler (2012). Creates overlay plots of the stretched and shifted LCs - batch 4.

   These scripts are called from within WHO5. An example of a plot created by these routines is given in Fig. L.2

2. **LCP3v_in:** Creates plots of the stretched and shifted LCs for each individual SN in V-band. The output is shown in Appendix A, each individual plot.

3. **LCP3v_pairOverlay:** Creates overlay plots of the stretched and shifted LCs for each individual pair of SNe in V-band. Commented out currently in WHO5, but can be used to investigate the LC pair further. An example of its output is shown in Fig. L.1.

4. **LCP3u:** Sadler (2012), creates overlay plots of the stretched and shifted LCs for u-band, not used in this thesis.

5. **ssptv:** Sadler (2012), creates differential plots (residuals and weighted signals) for V-band LCs. Examples of its output can be seen in Appendix B, left-hand side figures.

6. **ssptu:** Sadler (2012), creates differential plots (residuals and weighted signals) for u-band LCs, not used in this thesis.

7. **wscrp:** Sadler (2012), creates secondary parameter covariance plots. Examples of its output can be seen in Appendix B, right hand side figures.

8. **LCP5:** Sadler (2012), plots overlaid scaled PC signals.

9. **comparemmscdnfitWOerrpl:** In folder "subsamples". Plots the comparison between the physical secondary parameter values of SNe between the full sample and the subsamples. An example of the output is individual plots in Figs. 3.16 & 3.17.

10. **reddeningfreqpl:** Plots the reddening distribution histogram. In folder "REDDENING". Output is shown in the upper plot of Fig. 3.3.

11. **reddenningpl:** Plots the relationship between M_{MS} and ρ_c values and reddening of the SNe in the sample, marks the average M_{MS} and ρ_c values of the low, medium, and high reddening groups separately. In folder "REDDENING". Output is shown in the lower plot of Fig. 3.3.

12. **reddenning99aa91Tpl:** Same as "reddenningpl", but only for 99aa-likes and 91T-likes. In folder "REDDENING". Output shown in Fig. 4.3.

13. **mmscdncorrplSpiralNOerror:** Plots M_{MS} and ρ_c correlation for all SNe, marks normal-bright, underluminous, 91T-like, and 99aa-like objects separately. In folder "MMSCDNcorr". Output shown in Fig. 4.2.

14. **plHostdm15corrAll:** Plots the correlation between $\Delta m_{15}(B)$, s, and host-types for all SNe used in our sample, and the outliers. Normal-bright, 91bg-like, and outliers are marked separately. In folder "HostDm15SNtypeRelation_all". The output is shown in Fig. 4.7.

15. **mmsfreqLpl:** Plots M_{MS} distribution with normal-bright, transitional, and subluminous SNe marked with patterns. In folder "NEWerrplots". The outputs are shown in Figs. 4.1, 4.4, and 4.5 top plots.

16. **cdnfreqLpl:** Same as "mmsfreqLpl", but for ρ_c. In folder "NEWerrplots". The outputs are shown in Figs. 4.1, 4.4, and 4.5, bottom plots.

17. **mmsfreqYoungplot:** In folders "PlusMinusDwarfE", "PlusMinusDwarfEbulge", and "PlusMinusDwarfEbulgeS0". Plots the M_{MS} distribution for SNe in the spiral, S0, irregular, and small E hosts in folder "PlusMinusDwarfE". In folder "PlusMinusDwarfEbulge", this script excludes the SNe in spiral bulges, and in folder "PlusMinusDwarfEbulgeS0", it excludes SNe in S0 galaxies. Run FORTRAN code "plotml.f" (placed in folder "CODES") to prepare the data file required for the plotting routine to work. The make file for this code is "Makefile.plotml", placed in "SN-DATA-RUN" (the main directory containing WHO5). Run "make

-f Makefile.plotml" and then "xplotml" to run the code. Input files for the code that must be present in the main directory are "hostmass130.dat", "in.dat.opt", and "plotdataPrimary". The "hostmass130.dat" file contains the host type information for all SNe. "plotml.f" prepares "cdnfreqhistEllipticalBatchX.dat" or "cdnfreqhistEllipticalBatchX.dat" files, on which the distribution scripts can run. The outputs are shown in Figs. 4.8, 4.9, and 4.10, top left plots.For example, a part of the "hostmass130.dat" file is shown below:

AS15db	12.8824955	3	bulge
AS15dd	25.7039578	1	large E
AS15eb	91.2010839	2	S0
AS15ga	23.9883292	2	S0
AS15gr	0.87096359	4	spirals

Only the names of the SNe with six letters have been used in the first column. In the second column, the host type codes 1, 2, 3, and 4 have been used for large ellipticals, S0, spiral bulge, and spiral hosts, respectively. Code 5 is used when no host mass information is available. See Fig. 4.7 for the codes used in a plot.

18. **cdnfreqYoungplot:** In folders "PlusMinusDwarfE", "PlusMinusDwarfEbulge", and "PlusMinusDwarfEbulgeS0". Same as "mmsfreqYoungplot", but for ρ_c. Running instructions are the same as "mmsfreqYoungplot". The outputs are shown in Figs. 4.8, 4.9, and 4.10, top right plots.

19. **mmsfreqOldplot:** In folders "PlusMinusDwarfE", "PlusMinusDwarfEbulge", and "PlusMinusDwarfEbulgeS0". Same as "mmsfreqYoungplot", but for SNe in large E hosts in folder "PlusMinusDwarfE". In folder "PlusMinusDwarfEbulge", this script includes the SNe in spiral bulges, and in folder "PlusMinusDwarfEbulgeS0", it includes SNe in S0 galaxies. Running instructions are the same as "mmsfreqYoungplot". The outputs are shown in Figs. 4.8, 4.9, and 4.10, bottom left plots.

20. **cdnfreqOldplot:** In folders "PlusMinusDwarfE", "PlusMinusDwarfEbulge", and "PlusMinusDwarfEbulgeS0". Same as "mmsfreqOldplot", but for ρ_c. Running instructions are the same as "mmsfreqYoungplot". The outputs are shown in Figs. 4.8, 4.9, and 4.10, bottom right plots.

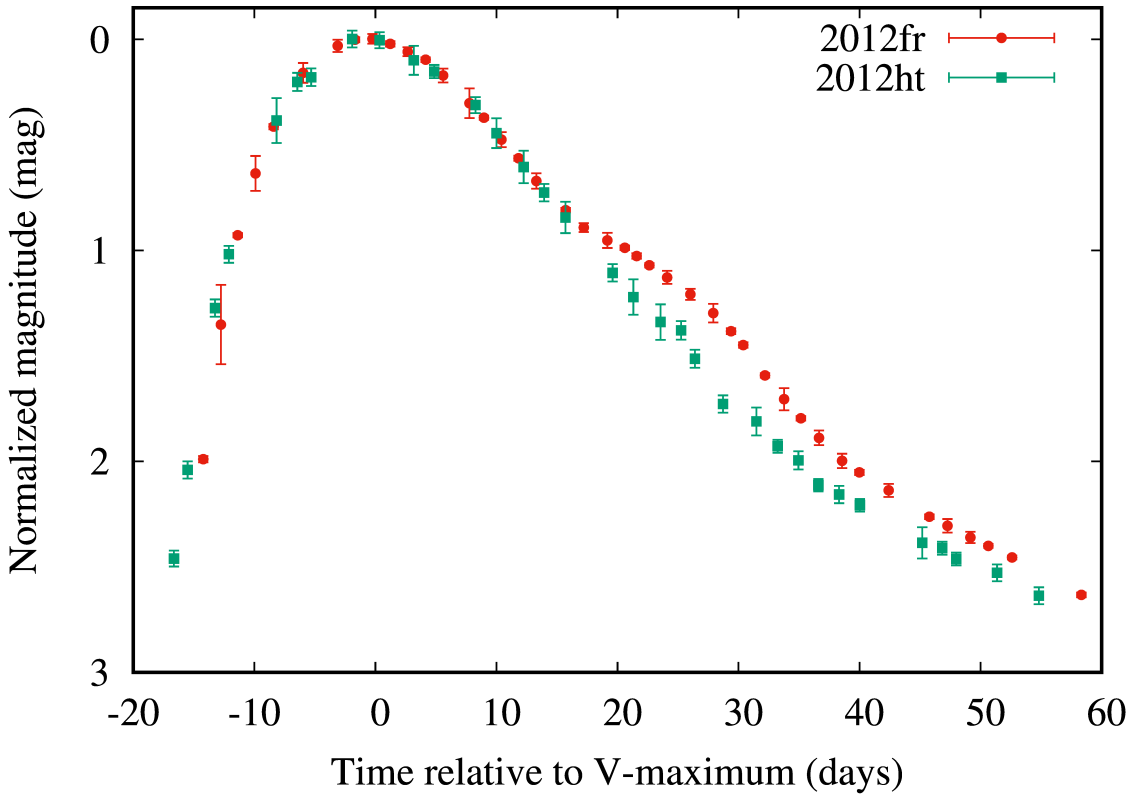

Figure L.1: Example of a stretched LC pair overlay plot, done with script "LCP3v_pairOverlay".

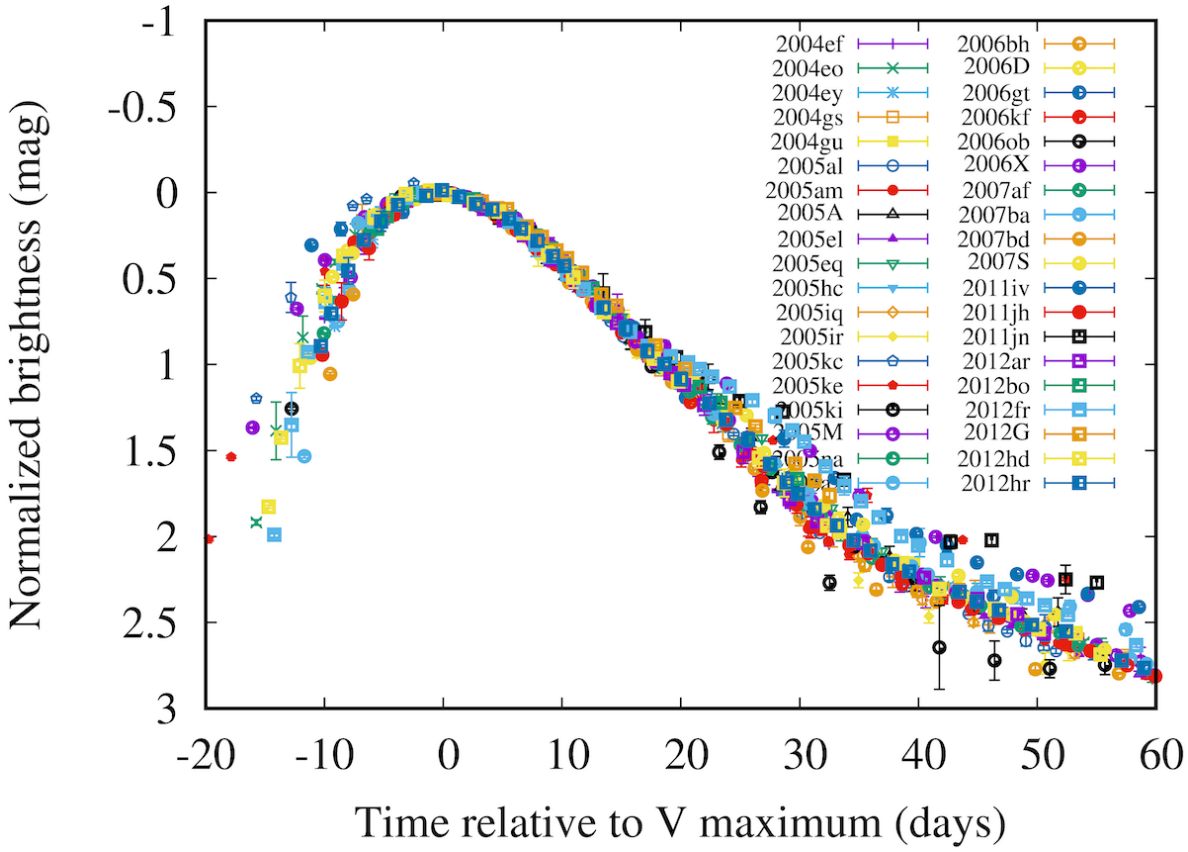

Figure L.2: Example of a stretched LC overlay plot, multiple LCs, done with script "LCP3v_d1".

L.3 WHO5 script

File "input.par" is created in WHO5 itself, so unlike Sadler (2012), there is no need to write this file manually anymore.

The user must create the following files in the main folder SN-DATA-RUN (same folder as the WHO5 script):

1. File "in.dat_all". Modified from Sadler (2012). This file should contain the names of the SNe in the sample. The names must be in the format of "123456.y.fin" - here the "123456" is the name or the abbreviation of the name of the SN, and it should contain a combination of six letters, numbers, and/or special characters. The "y" is the band-name (e.g. for V-band, use "v", for B-band, use "b", and for u-band, use "u".), and the last three characters must be "fin". The first few names (presently eight, can be modified by the future user) of the "in.dat_all" file must be the "messenger" SNe Ia, these eight SNe must be present in all runs. Example of "in.dat_all" file:

2004ef.v.fin
2004gu.v.fin
2005iq.v.fin
2005ke.v.fin
2005na.v.fin
2006ax.v.fin
2006gt.v.fin
2006X_.v.fin
2004eo.v.fin
2004ey.v.fin
2004gs.v.fin
2005A_.v.fin
2005al.v.fin
2005am.v.fin
2005el.v.fin
2005hc.v.fin
2005kc.v.fin
2005eq.v.fin
2005ir.v.fin
2005ki.v.fin
2005M_.v.fin
2006bh.v.fin
2006D_.v.fin
2006kf.v.fin
2006ob.v.fin
2007S_.v.fin
2007af.v.fin
2007ba.v.fin
2007bd.v.fin
2011iv.v.fin
2011jh.v.fin
2011jn.v.fin

Only one column of data here gives the SNe names used in the analysis. Notice that the first eight names are of the "messenger" SNe Ia used in our analysis. The first few names must always be the messenger SNe. Here the CSP-I names were placed first, and then the CSP-II names, though this is by choice, and not necessary.

2. File "redshifts25.dat_all". Modified from Sadler (2012). This must have the names and the redshifts (in floating point) of the same SNe names mentioned in "in.dat_all". The names may or may not be in the same order as "in.dat_all". Example of "redshifts25.dat_all" file:

2004ef.v.raw 0.029771
2004gu.v.raw 0.046897
2005iq.v.raw 0.032929
2005ke.v.raw 0.004483
2005na.v.raw 0.026815
2006ax.v.raw 0.017957
2006gt.v.raw 0.043641
2006X_.v.raw 0.006322
2004eo.v.raw 0.014734
2004ey.v.raw 0.014627
2004gs.v.raw 0.027498
2005A_.v.raw 0.018340
2005al.v.raw 0.013286
2005am.v.raw 0.008967
2005el.v.raw 0.014885
2005hc.v.raw 0.044983
2005kc.v.raw 0.013890
2005eq.v.raw 0.028351
2005ir.v.raw 0.075254
2005ki.v.raw 0.020370
2005M_.v.raw 0.022972
2006bh.v.raw 0.010493
2006D_.v.raw 0.009641
2006kf.v.raw 0.020799

The names of the SNe are in the first column and the corresponding redshifts are in the second column. The redshift values all have six digits after the decimal.

In the next step, the system is run with the following command (modified from Sadler 2012.):

WHO5 name color PCchoice

name: this is user-defined and arbitrary.

color: band analyzed - v, b, or u.

PCchoice: type of PC analysis - PCd (differential) or PCr (residual) analysis.

Example of running command:

WHO5 new v PCd

The name of the run here is "new". It is a V-band differential analysis.

A guideline of how WHO5 works:

1. WHO5 runs "DATACHECK_NEW1", which checks if there is LC data in region II, and writes "in.dat" and "redshifts25.dat" files.

2. "NEWCHECK" removes duplicate LC data points.

3. Primary parameter file "in.dat.opt" is searched for (Sadler 2012).

4. If "in.dat.opt" is not present, then ATEST1_tcsh is run, which calls a subscript named ATEST, and creates "in.dat.opt". "NARC" looks for and fills the large gaps in LC data with template points. ATEST is called again and the primary parameters are calculated from the augmented LCs, and "in.dat.opt" is modified with them (Sadler 2012).

5. "DATACHECK_NEW2" checks if LC data covers up to at least +25 days, if not, then discards the SNe, and rewrites "in.dat" and "redshifts25.dat" files.

6. Next, SADIE013 modifies the "in.dat.opt" and saves the old "in.dat.opt" as "in.dat.opt.1" (Sadler 2012).

7. ALPHA11 reads "in.dat.opt", as well as LC data from the RDAT folder, and stretches the LC data (Sadler 2012). The stretched LC data are written in the temporary folder named "CDAT". In the process of doing this, this code also does the correction due to the brightness of the objects, by calling a subroutine named "qcorr.f". See Chapter 3.1.2 for the details of the correction.

8. For B or V bands, "LCP3v_d1", "LCP3v_d2", "LCP3v_d3", "LCP3v_d4" plotting scripts are called. They create LC overlay plots by breaking the sample into four groups (for a large number of SNe, the overlay is hard to see if all LCs are overlaid together in one single plot). "LCP3v_in" is called next. This script plots the individual stretched and shifted LCs separately. In the case of u-band, these are replaced with "LCP3u".

9. The names of the stretched and shifted LC data from the "CDAT" folder are written in a file called "sadie.dat". WHO5 looks for the signals of the color analyzed in the folder "SDAT". Same as Sadler (2012).

10. Next, the presence of the secondary signals is determined by "SADIE". Differentials and residuals are created and stored in folders named "DIFL" and "RSDL", respectively. Parameter covariance data is stored in the folder "GRID". Same as Sadler (2012).

11. The differentials and residuals are plotted next ("ssptv" script for B and V-bands, "ssptu" for u-band). Input data folders are "DIFL", "RSDL", and "SSIG". Same as Sadler (2012). Later these plots will be placed in the folder "nameRES" where "name" is the user-defined run name.

12. Signal coefficient covariance plots are created by "wscrp". Same as Sadler (2012). Later these plots will be placed in the folder named "nameTST". Like before, "name" is the run name.

13. For u-band, UDEV creates files named udev0.dat, udev1.dat, and udev2.dat (Sadler 2012). This script is not used in this thesis, as the only band this work handles is V.

14. Next, the codes in the folder named "Peter" are run. They translate the generic progenitor parameters into physical secondary parameters by using theoretical relationships. Also, Monte Carlo is done to determine the possible physical range covered by the limited number of SNe in the sample. Important things to note:

    - If using only the CSP-I SNe, then it should be a "reference" run.
    - A file called "switch.dat" must be present in the folder Peter, this must have two integers, "switch" and "jump". The only accepted values for these parameters are 0 or 1. "switch" should be 1 for a reference run (CSP-I), and 1 for all non-reference runs (mix of CSP-I and II, only CSP-II, or any other data, except for just CSP-I). The parameter jump should always be 1 unless the Monte-Carlo method is wished to be avoided.
    - A file named "plotdataPrimaryBen" should be present, this is for use in the reference run only. This file should not be modified, it contains the generic and physical secondary parameter values from Sadler (2012).
    - If it is a reference run, then the generic and physical secondary parameter values from the run are stored in a file called "plotdataPrimaryBig", used as a standard in all non-reference runs.
    - The "messenger" SNe names are stored in a file called "names_CSPI8.dat". This file is used for all runs.
    - The results of this part of WHO5 are stored in files named "order.dat" and "plotdataPrimary". This "plotdataPrimary" file would be required for plotting all the secondary parameter distributions later.

15. PC analyzer is initiated by "EPSILv" (for B and V-bands) or "EPSILu" (for u-band). Same as Sadler (2012).

16. The overdetermined PC coefficient system is solved by "EPCSOLV4v" (B or V-band differential analysis) or "EPCSOLV3v" (B or V-band residual analysis), or "EPCSOLV4u" (u-band differential analysis), or "EPCSOLV3u" (u-band residual analysis) script. All differentials or residuals are plotted and their optimal scaled PCs are overlaid. An overlay of all scaled PCS is plotted by LCP5. Same as Sadler (2012).

17. **Clean up:** Temporary files are deleted, or stored if necessary (examples of stored files are ".sm1", and ".sm2" files, which can be used for re-plotting the χ^2 maps later, if any changes are necessary. This can be easily switched off in WHO5). The log of the run, input and output data, and the results along with the source codes are packaged in a tar file and then placed in the folder "RUNS". Modified from Sadler (2012).

18. **"epstojpg_all"** command, where used, converts the ".eps" files to ".png". This command is switched off inside WHO5 currently, as for our $\sim 23,000$ plots it takes nearly two days to convert all of them to ".png" format. Therefore all image files are stored in ".eps" format now. The "epstojpg_all" command can be run over all or a subset of plots at any time, or, if wished, can be switched on inside WHO5, so that the conversion will be on.

19. The run folder is placed in the main directory. **WHO5 is complete.**

20. **Finding the results in the run folder:**

    - The results of a run are stored in a directory called "RUNname", produced by the run, where "name" is the user-given run name. Same as Sadler (2012).
    - In the "RUNname" directory, the "DATA" folder contains copies of raw data in "RDAT", stretched data in "CDAT", templates in "TDAT", and signals in "SDAT". Same as Sadler (2012). Modification in this work: the run directory also contains a sub-folder named "DATA_TST" which has the necessary files for plotting the signal coefficient covariance plots (just in case any modifications are needed in the plots).
    - In "RUNname", folder "nameRES" contains all differential plots and scaled signals. "name" is the run name.
    - "nameTST" folder contains signal coefficient covariance plots. "name" is the run name.
    - The folder "LC" contains all individual and overlaid stretched LC plots.
    - "runlog.name" file contains a detailed log of the run.
    - Folder "UNCERTAINTY" contains files containing information to determine t_{cor} and f_{cor} (equation 3.2) via Fig. 3.9 if needed, which is in the case of a new data set being used.
    - "DPAR" folder contains files "cof.dat", "epccof.dat", "primary.dat", "epc.dat", and "avg.dat". Same as Sadler (2012).
    - Folder "EPJPG" contains differentials or residuals with the optimally scaled PC signal overlaid. Same as Sadler (2012).
    - folder "PPAR" contains all header files and files "temp.dat" and "sig.dat". Same as Sadler (2012).
    - Folder "SRC" contains copies of all source codes and scripts, along with WHO5. Same as Sadler (2012).