

**NAME**

scans.4 – read data from ASCII scan files

**SYNOPSIS**

```
fn scans.4 [options scan_numbers]
```

**DESCRIPTION**

The `scans.4` function reads in files of ASCII data according to a modest set of conventions.

When used with X-ray scattering data, `scans.4` can perform scan averaging, background subtraction, data normalization and error-bar calculation. However, `scans.4` works well with any kind of data file that follows the conventions described below.

**COMMAND LINE OPTIONS**

```
.           Use same options as last time
-i          Initialize, used to start function and return
-f file    Select scan file name
-p          Print scan-file contents
+d or -d    Collect (or don't) 3 columns of data
+e or -e    Calculate (or don't) error bars from statistics
+M or -M    Use (or don't) special MCA data convention
+n or -n    Normalize (or don't) data points
+o or -o    Sort (or don't) data points
+q or -q    Don't (or do) print messages (quiet)
+r or -r    Rerange (or don't) plot axes for each new data set
+s or -s    Sort and merge (or don't) data by x values
+v or -v    Print (or don't) each line of scan file (verbose)
+I or -I    Use (or don't) #I intensity normalization
+S or -S    Retrieve scans by scan (or file) number
x=#         Set column for x values
y=#         Set column for y values
z=#         Set column for z values, turn on +d flag
m=#         Set col for monitor normalization, turn on +n flag
t=#         Set col for time normalization, turn on +n flag
x=M         Stuff MCA channel numbers in x in 3D mode
y=M         Stuff MCA channel numbers in y in 3D mode
```

The default settings correspond to the following options in 2D C-PLOT mode

```
-f data +eosSn -drvIqM x=1 y=-1
```

and

```
-f data +deosSn -rvIqM x=1 y=2 z=-1 x=M
```

in 3D C-PLOT mode.

**SPECIFYING SCANS**

Scans can be retrieved by entering either the scan number (option `+S`, the default) or the file position number (option `-S`). Scan numbers are determined by the `#S` lines in the file. The file position number is the sequence position of the scan in the file, irrespective of scan number.

When selecting by scan numbers, if there is more than one scan with the same number, the last of them is retrieved. You can specify which instance of a repeated scan number to retrieve by appending a decimal point and an index number to the scan number. For example, selecting scan number `10.3` retrieves the third scan from the start of the file that has

scan number 10.

Negative numbers count back from the end of the file and are always considered to be file-position numbers. For example,

```
fn . -1
```

will always return the last scan in the file.

You can enter multiple scan numbers to select the scans you are interested in. Scan numbers that end with `b` are used as background scans. For example,

```
fn . 12b 13 14b 15b 16 17b
```

Data in the background scans will be subtracted from the data in the non-background scans that has corresponding  $x$  values. Choosing a background scan will force the data to be sorted by  $x$  values.

You can read in a group of consecutive scans with

```
fn . 3-7 10-14
```

This command would read in scans 3 through 7 and 10 through 14.

## FILE CONVENTIONS

The scan files contain control lines, data lines and blank lines. Control lines contain a `#` character in the first column followed by a command word. Data lines generally contain a row of numbers. Special data lines containing MCA data begin with an `@` character followed by a row of numbers. These data lines are ignored unless the use MCA data option `+M` is selected.

The control conventions used by *scans.4* are as follows:

`#S N` – starts a new scan. Here,  $N$  is the user's numbering scheme and is the number used when retrieving by scan number (`+S`). Most often the scan number is the position of the scan in the file.

`#M N` – indicates data was taken counting to *number* monitor counts.

`#T N` – indicates data was taken counting for  $N$  seconds.

`#N N [M]` – indicates there are  $N$  columns of data. If  $M$  is present, it indicates there are  $M$  sets of data columns on each line. When collecting data from a multi-channel analyzer, for example, the data might be arranged with 16 points per line in the file to make the file easier to scan by eye. In such a case, the control line would be `#N 1 16`.

`#I N` – is for an optional multiplicative intensity-normalization factor.

`#@MCA` – indicates the scan contains MCA data. If the `+M` option is selected,  $x$  (2D or 3D) or  $y$  (3D only) values will be calculated automatically. In three-column mode, whether it is  $x$  or  $y$  depends on whether the `x=M` or `y=M` command line option is selected or on which interactive response was given. Data in the lines starting with `@A` will be stuffed into the  $y$  (2D) or  $z$  (3D) data array.

`#@CALIB a b c` – gives calibration factors for MCA data. The  $x$  (2D or 3D) or  $y$  (3D only) values will be calculated using the formula

$$x[i] = a + b * i + c * i * i$$

where  $i$  is the point number, starting from zero. Calibration factors can be changed within the data portion of a scan for subsequent MCA data by the line

```
@CALIB a b c
```

Before each scan is read by *scans.4*, the calibration parameters are initialized to zero.

The following control lines are not commands but are printed out as they are encountered while reading a scan:

#C – is a comment line.

#D – is followed by the date and time the scan was taken.

#L *label1 label2* – is the data-column labels, with each label separated from the next by *two* spaces.

For example, a very simple file might have:

```
#S 1
#N 3
#L Temperature Voltage Counts
23.4 1.01 30456
23.6 1.015 24000

#S 2
etc.
```

## DATA COLUMNS

When running C-PLOT in 2D mode, the default behavior is to take  $x$  values from the first column and  $y$  values from the last column. If in 3D mode,  $x$  values are taken from the first column,  $y$  values from the second and  $z$  values from the last column. If normalizing the data, the default behavior is to use the #T or #M values. If neither appear, you must enter a column number for the normalization values.

When entering column numbers, a negative number counts backward from the last column. If the column for  $x$  is zero, the value put in for  $x$  is just the index number of the point.

## ENTERING OPTIONS

If you give a dot `.` as the command-line argument or in response to “Scans/options”, the same argument or option string will be used as last time. That is, the string is remembered, not the options chosen interactively using “Change modes?”. For instance, if you enter a long sequence of scan numbers and read in the scans, then change something via “Change modes?”, you can simply enter a dot in response to “Scans/options” and recover the previous sequence of scan numbers.

When you do enter a string of flags and scan numbers, the modes set by the flags only apply to the scans that follow the flags, not the preceding scans.

## THE INDEX FILE

Reading a long ASCII data file takes time. When `scans.4` first opens a file, it scans the whole file and saves a directory of the scans in a binary-format index file. The name of the index file is formed by appending `.I` to the original data-file name.

As long as the index file is more recent than the data file, `scans.4` will use the information in the index file.

## NORMALIZATION AND ERROR BARS

Data can be normalized to either monitor counts or time. When normalizing to monitor counts, the error bars will include the uncertainty in the counting statistics of the monitor counts. Otherwise there is no difference between specifying time or monitor counts.

By default, `scans.4` normalizes data to monitor counts, with the second to last data column used for the monitor count values. Use the `-n` flag to turn off normalization. If a column number is selected using the `m=col` or `t=col` arguments, normalization is set to monitor or

time mode, respectively, using the column number specified. If the column number in either case is given as zero, the normalization mode and value given by the #M or #T directives for a particular scan in the data file are used. It is an error for normalization mode to be on, for the normalization column to be set to zero and for no normalization directives to be present for a scan.

The normalization modes selected remain in effect for subsequent scans.

The values returned as error bars are those due to counting statistics (the square root of the number of counts). When the counts are derived from the algebraic combination of detector, background and monitor counts, the error bars are calculated using the appropriate "propagation of errors" formalism. See the source code for details.

If the +I option is selected, the counts for each point are multiplied by the value given by the #I control line in the scan header. If the +I option is selected and the scan header doesn't contain a #I control line, the counts are not changed.

#### NUMBER OF POINTS AND SCANS

Earlier version of *scans.4.c* had built-in limits to the number of scans or raw data points that could be handled. Those limits no longer exist.