

Open Raw File Description

Kyriacos Leptos, Dat Nguyen,
September 2003

0. Introduction

The raw data of an experiment are stored in three different types of directories within a folder named after it (fig. 1).

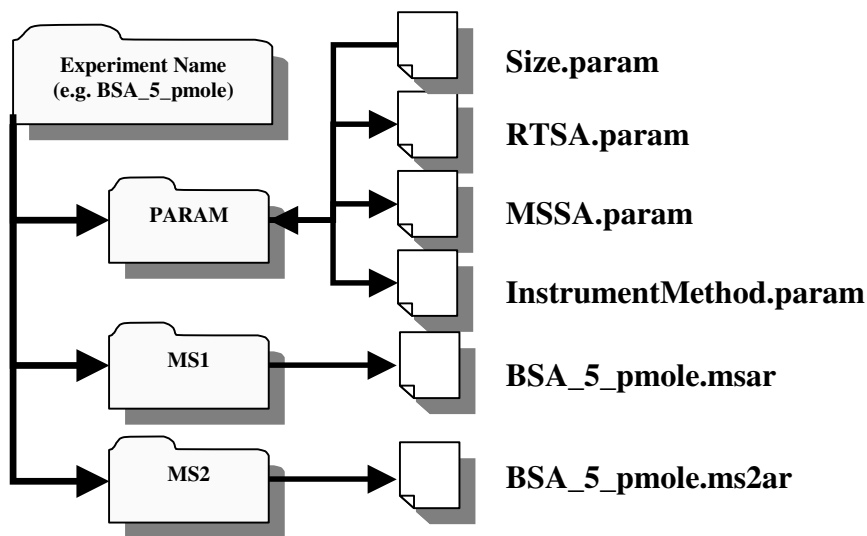


Fig 1. The OpenRaw format (file tree structure)

These directories are:

- a. The global parameters directory
- b. The MS1 spectra archive directory
- c. The MS2 spectra archive directory

1. The Global Parameters Directory

There are four text files stored in this directory. The *size.param* file stores information about the size of the data of an LC/MS experiment, i.e. total number of scans and total number of mass bins. The *RTSA.param* file, which stands for *Retention Time Sampling Array*, stores information about the time points at which each mass spectrum was scanned. The *MSSA.param* file, which stands for Mass Sampling Array, stores information about the spacing of the sampling points in the m/z dimension given by the spectrometer. This file is produced only if the data is acquired in the profile mode. The file *InstrumentMethod.param* stores information about the instrument method used by the user at that particular LC/MS experiment.

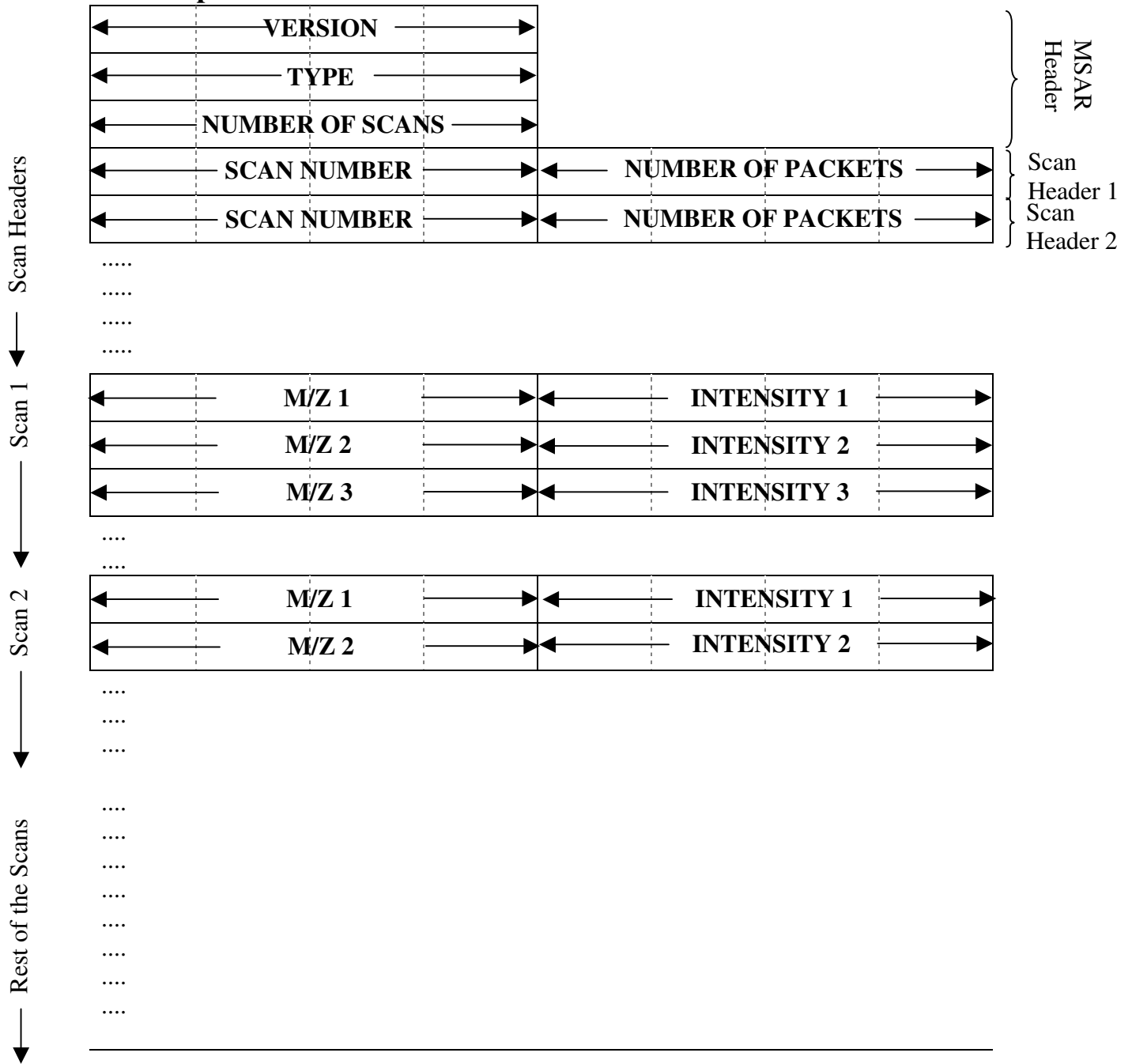
2. The MS1 Spectra Directory

In this directory the binary file *expmnt_name.msar* is stored. The extension *msar* stands for mass-spectrum (MS) archive. The name archive indicates that the ion-abundance signal is stored in the format of an array of mass spectra. Each mass spectrum is itself in the format of an array of ion-abundance for each m/z data point (see further for *msar* description).

3. The MS2 Spectra Directory

This directory holds the binary file *expmnt_name.ms2ar*. The extension *ms2ar* stands for MS/MS archive and its format is similar to the *msar* archive mentioned above. (see further for *ms2ar* description).

4. Description of MSAR file



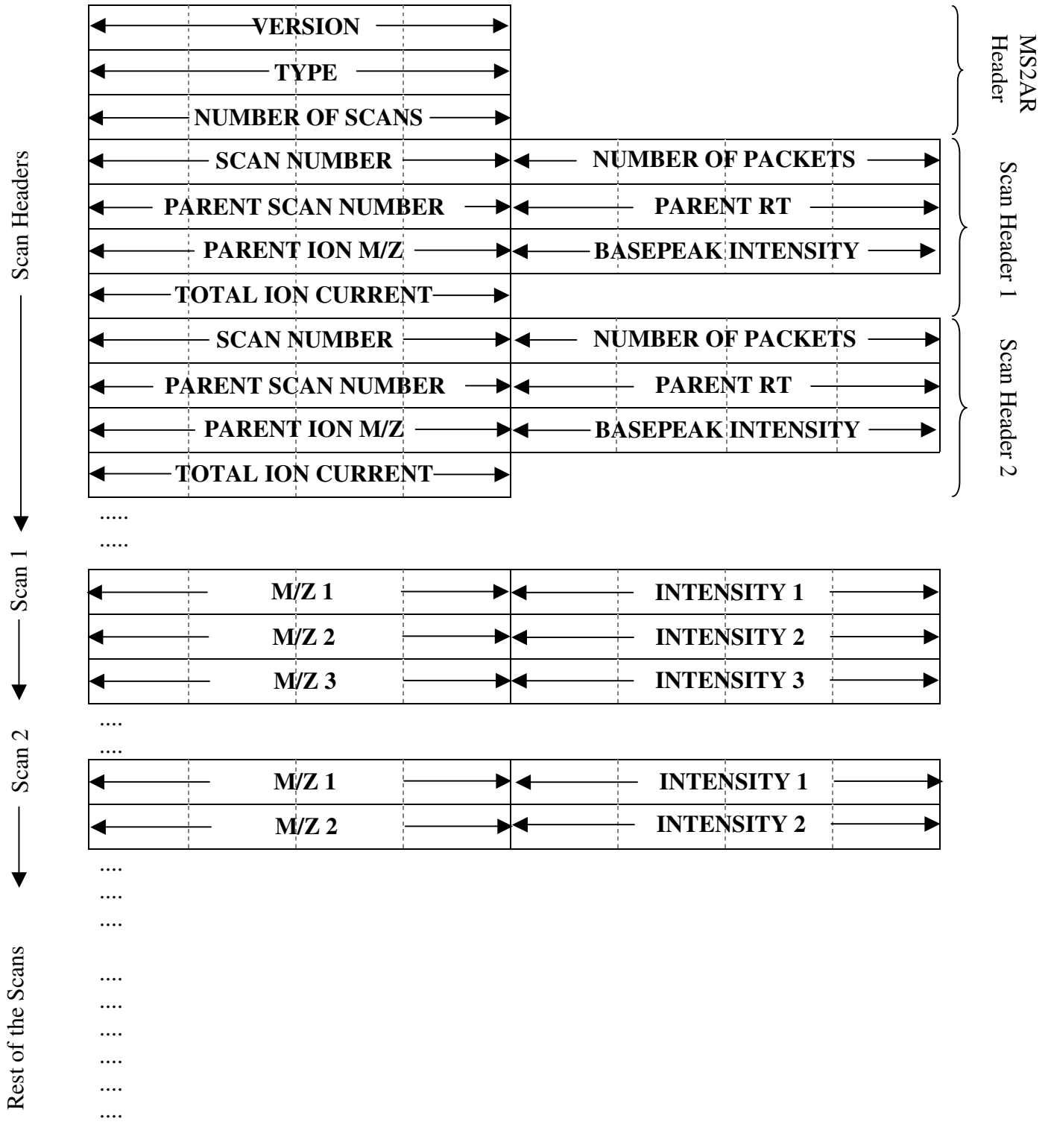
MSAR Header Byte Description

Bytes	Type	Description
0...3	float	Version
4...7	int	Type 0 = PROFILE, 1 = CENTROID If type is equal to PROFILE then the packets for each scan are 4 bytes long otherwise they are 8 bytes long as shown above
8...11	int	Number of scan headers to follow. A scan header is composed of two integer values; the number of the scan and the number of packets for than scan.

Scan Header Byte Description

Bytes	Type	Description
0...3	int	Scan number
4...7	int	Number of packets

5. Description of MS2AR file



MS2AR Header Byte Description

Bytes	Type	Description
0...3	float	Version
4...7	int	Type 0 = PROFILE, 1 = CENTROID If type is equal to PROFILE then the packets for each scan are 4 bytes long otherwise they are 8 bytes long as shown above
8...11	int	Number of scan headers to follow. A scan header is composed of two integer values; the number of the scan and the number of packets for than scan.

Scan Header Byte Description

Bytes	Type	Description
0...3	int	Scan number
4...7	int	Number of packets
8...11	int	Number of parent scan
12..15	float	Retention time of parent scan
16..19	float	M/Z of parent ion
20..23	float	Basepeak, i.e. the most abundant peak in the spectrum
24..27	float	Total Ion Current (TIC)