# asammdf Documentation

Release 2.3.2

**Daniel Hrisca** 

## Contents

1	Project goals	3
2	Features	5
3	Major features still not implemented	7
4	Dependencies	9
5	Installation	11
6	API	13
7	Benchmarks	47
8	Indices and tables	57

asammdf is a fast parser/editor for ASAM (Associtation for Standardisation of Automation and Measuring Systems) MDF (Measurement Data Format) files.

asammdf supports both MDF version 3 and 4 formats.

asammdf works on Python 2.7, and Python >= 3.4

Contents 1

2 Contents

Project goals

The main goals for this library are:

- to be faster than the other Python based mdf libraries
- to have clean and easy to understand code base

## **Features**

- read sorted and unsorted MDF v3 and v4 files
- files are loaded in RAM for fast operations
  - for low memory computers or for large data files there is the option to load only the metadata and leave the raw channel data (the samples) unread; this of course will mean slower channel data access speed
- extract channel data, master channel and extra channel information as *Signal* objects for unified operations with v3 and v4 files
- time domain operation using the Signal class
  - Pandas data frames are good if all the channels have the same time based
  - usually a measuremetn will have channels from different sources at different rates
  - the Signal class facilitates operations with such channels
- remove data group by index or by specifing a channel name inside the target data group
- · append new channels
- filter a subset of channels from original mdf file
- convert to different mdf version
- add and extract attachments
- mdf 4.10 zipped blocks

6 Chapter 2. Features

## Major features still not implemented

- functionality related to sample reduction block (but the class is defined)
- mdf 3 channel dependency save and append (only reading is implemented)
- handling of unfinnished measurements (mdf 4)
- mdf 4 channel arrays
- xml schema for TXBLOCK and MDBLOCK

## Dependencies

## asammdf uses the following libraries

• numpy : the heart that makes all tick

• numexpr : for algebraic and rational channel conversions

matplotlib : for Signal plottingpandas : for DataFrame export

## Installation

## asammdf is available on

• github: https://github.com/danielhrisca/asammdf/

• PyPI: https://pypi.org/project/asammdf/

pip install asammdf

API

## **MDF**

This class acts as a proxy for the MDF3 and MDF4 classes. All attribute access is delegated to the underling *file* attribute (MDF3 or MDF4 object). See MDF3 and MDF4 for available extra methods.

**class** a sammdf.mdf.mdf (name=None, load\_measured\_data=True, version='3.20') Unified access to MDF v3 and v4 files.

Parameters name: string

mdf file name

load\_measured\_data : bool

load data option; default True

- if True the data group binary data block will be loaded in RAM
- if False the channel data is read from disk on request

version: string

mdf file version ('3.00', '3.10', '3.20', '3.30', '4.00', '4.10', '4.11'); default '3.20'

## Methods

convert
filter
iter\_to\_pandas

#### convert (to)

convert MDF to other versions

Parameters to: str

```
new mdf version from ('3.00', '3.10', '3.20', '3.30', '4.00', '4.10', '4.11')
```

**Returns out**: MDF

new MDF object

filter(channels)

return new MDF object that contains only the channels listed in channels argument

Parameters channels: list

list of channel names to be filtered

Returns mdf: MDF

new MDF file

iter\_to\_pandas()

generator that yields channel groups as pandas DataFrames

## MDF3 and MDF4 classes

#### MDF3

asammdf tries to emulate the mdf structure using Python builtin data types.

The *header* attibute is an OrderedDict that holds the file metadata.

The groups attribute is a dictionary list with the following keys:

- data\_group : DataGroup object
- channel\_group : ChannelGroup object
- channels: list of Channel objects with the same order as found in the mdf file
- channel\_conversions : list of ChannelConversion objects in 1-to-1 relation with the channel list
- channel\_sources : list of SourceInformation objects in 1-to-1 relation with the channels list
- data\_block : DataBlock object
- texts : dictionay containing TextBlock objects used throughout the mdf
  - channels: list of dictionaries that contain TextBlock objects ralated to each channel
    - \* long\_name\_addr : channel long name
    - \* comment\_addr : channel comment
    - \* display\_name\_addr : channel display name
  - channel group: list of dictionaries that contain TextBlock objects ralated to each channel group
    - \* comment\_addr : channel group comment
  - conversion\_tab: list of dictionaries that contain TextBlock objects ralated to VATB and VTABR channel conversions
    - \* text\_{n}: n-th text of the VTABR conversion

The *file\_history* attribute is a TextBlock object.

The *channel\_db* attibute is a dictionary that holds the (*data group index, channel index*) pair for all signals. This is used to speed up the *get\_signal\_by\_name* method.

The *master\_db* attibute is a dictionary that holds the *channel index* of the master channel for all data groups. This is used to speed up the *get\_signal\_by\_name* method.

## **API**

class asammdf.mdf3.MDF3 (name=None, load\_measured\_data=True, version='3.20')

If the name exist it will be loaded otherwise an empty file will be created that can be later saved to disk

Parameters name: string

mdf file name

load\_measured\_data : bool

load data option; default True

- if True the data group binary data block will be loaded in RAM
- if False the channel data is read from disk on request

version: string

mdf file version ('3.00', '3.10', '3.20' or '3.30'); default '3.20'

## **Attributes**

name	(string) mdf file name
groups	(list) list of data groups
header	(OrderedDict) mdf file header
file_history	(TextBlock) file history text block; can be None
load_measured_d	atabool) load measured data option
version	(int) mdf version
channels_db	(dict) used for fast channel access by name; for each name key the value is a (group
	index, channel index) tuple
masters_db	(dict) used for fast master channel access; for each group index key the value is the
	master channel index

## Methods

add_trigger
append
get
info
iter_get_triggers
remove
save

```
trigger time

pre_time: float

trigger pre time; default 0

post_time: float

trigger post time; default 0

comment: str

trigger comment

append (signals, acquisition_info='Python', common_timebase=False)

Appends a new data group.

Parameters signals: list

list on Signal objects

acquisition_info: str

acquisition_info: str

acquisition information; default 'Python'

common_timebase: bool

flag to hint that the signals have the same timebase
```

## **Examples**

```
>>> # case 1 conversion type None
>>> s1 = np.array([1, 2, 3, 4, 5])
>>> s2 = np.array([-1, -2, -3, -4, -5])
>>> s3 = np.array([0.1, 0.04, 0.09, 0.16, 0.25])
>>> t = np.array([0.001, 0.002, 0.003, 0.004, 0.005])
>>> names = ['Positive', 'Negative', 'Float']
>>> units = ['+', '-', '.f']
>>> info = {}
>>> s1 = Signal(samples=s1, timstamps=t, unit='+', name='Positive')
>>> s2 = Signal(samples=s2, timstamps=t, unit='-', name='Negative')
>>> s3 = Signal(samples=s3, timstamps=t, unit='flts', name='Floats')
>>> mdf = MDF3('new.mdf')
>>> mdf.append([s1, s2, s3], 'created by asammdf v1.1.0')
>>> # case 2: VTAB conversions from channels inside another file
>>> mdf1 = MDF3('in.mdf')
>>> ch1 = mdf1.get("Channel1_VTAB")
>>> ch2 = mdf1.get("Channel2_VTABR")
>>> sigs = [ch1, ch2]
>>> mdf2 = MDF3('out.mdf')
>>> mdf2.append(sigs, 'created by asammdf v1.1.0')
```

**get** (name=None, group=None, index=None, raster=None, samples\_only=False) Gets channel samples. Channel can be specified in two ways:

•using the first positional argument name

- -if there are multiple occurances for this channel then the *group* argument can be used to select a specific group.
- -if there are multiple occurances for this channel and the *group* argument is None then a warning is issued

•using the group number (keyword argument *group*) and the channel number (keyword argument *in-dex*). Use *info* method for group and channel numbers

If the raster keyword argument is not None the output is interpolated accordingly

```
Parameters name: string
          name of channel
        group: int
          0-based group index
        index: int
          0-based channel index
        raster: float
          time raster in seconds
        samples_only: bool
          if True return only the channel samples as numpy array; if False return a Signal object
    Returns res: (numpy.array | Signal)
          returns Signal if samples_only*=*False (default option), otherwise returns numpy.array
    Raises MdfError:
        * if the channel name is not found
        * if the group index is out of range
        * if the channel index is out of range
get MDF information as a dict
```

## **Examples**

info()

```
>>> mdf = MDF3('test.mdf')
>>> mdf.info()
```

## iter\_get\_triggers()

generator that yields triggers

## Returns trigger\_info: dict

trigger information with the following keys:

· comment : trigger comment

• time : trigger time

pre\_time : trigger pre timepost\_time : trigger post time

• index : trigger index

• group : data group index of trigger

```
remove (group=None, name=None)
```

Remove data group. Use *group* or *name* keyword arguments to identify the group's index. *group* has priority

Parameters name: string

name of the channel inside the data group to be removed

group: int

data group index to be removed

## **Examples**

```
>>> mdf = MDF3('test.mdf')
>>> mdf.remove(group=3)
>>> mdf.remove(name='VehicleSpeed')
```

**save** (*dst=None*)

Save MDF to dst. If dst is None the original file is overwritten

## MDF version 3 blocks

The following classes implement different MDF version3 blocks.

#### **Channel Class**

```
class asammdf.mdf3.Channel(**kargs)
```

CNBLOCK class derived from dict

The Channel object can be created in two modes:

- •using the file\_stream and address keyword parameters when reading from file
- •using any of the following presented keys when creating a new Channel

The keys have the following meaning:

- •id Block type identifier, always "CN"
- •block\_len Block size of this block in bytes (entire CNBLOCK)
- •next ch addr Pointer to next channel block (CNBLOCK) of this channel group (NIL allowed)
- •conversion\_addr Pointer to the conversion formula (CCBLOCK) of this signal (NIL allowed)
- •source\_depend\_addr Pointer to the source-depending extensions (CEBLOCK) of this signal (NIL allowed)
- •ch\_depend\_addr Pointer to the dependency block (CDBLOCK) of this signal (NIL allowed)
- •comment\_addr Pointer to the channel comment (TXBLOCK) of this signal (NIL allowed)
- •channel\_type Channel type
  - -0 = data channel
  - -1 = time channel for all signals of this group (in each channel group, exactly one channel must be defined as time channel) The time stamps recording in a time channel are always relative to the start time of the measurement defined in HDBLOCK.

- •short\_name Short signal name, i.e. the first 31 characters of the ASAM-MCD name of the signal (end of text should be indicated by 0)
- •description Signal description (end of text should be indicated by 0)
- •start\_offset Start offset in bits to determine the first bit of the signal in the data record. The start offset N is divided into two parts: a "Byte offset" (= N div 8) and a "Bit offset" (= N mod 8). The channel block can define an "additional Byte offset" (see below) which must be added to the Byte offset.
- •bit count Number of bits used to encode the value of this signal in a data record
- •data\_type Signal data type
- •range\_flag Value range valid flag
- •min\_raw\_value Minimum signal value that occurred for this signal (raw value)
- •max\_raw\_value Maximum signal value that occurred for this signal (raw value)
- •sampling\_rate Sampling rate for a virtual time channel. Unit [s]
- •long\_name\_addr Pointer to TXBLOCK that contains the ASAM-MCD long signal name
- •display\_name\_addr Pointer to TXBLOCK that contains the signal's display name (NIL allowed)
- •aditional\_byte\_offset Additional Byte offset of the signal in the data record (default value: 0).

## Parameters file\_stream : file handle

mdf file handle

address: int

block address inside mdf file

## **Examples**

## **Attributes**

name	(str) full channel name
address	(int) block address inside mdf file
dependencies	(list) lsit of channel dependencies

## **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
	Continued on next page

Table 6.3 – continued from previous page

fromkeys			
get items			
items			
keys			
pop			
popitem setdefault			
setdefault			
update values			
values			

#### ChannelConversion Class

```
class asammdf.mdf3.ChannelConversion(**kargs)
```

CCBLOCK class derived from dict

The ChannelConversion object can be created in two modes:

- •using the file\_stream and address keyword parameters when reading from file
- •using any of the following presented keys when creating a new ChannelConversion

The first keys are common for all conversion types, and are followed by conversion specific keys. The keys have the following meaning:

```
•common keys
```

```
-id - Block type identifier, always "CC"
```

- -block\_len Block size of this block in bytes (entire CCBLOCK)
- -range\_flag Physical value range valid flag:
- -min\_phy\_value Minimum physical signal value that occurred for this signal
- -max\_phy\_value Maximum physical signal value that occurred for this signal
- -unit Physical unit (string should be terminated with 0)
- -conversion\_type Conversion type (formula identifier)
- -ref\_param\_nr Size information about additional conversion data

## •specific keys

-linear conversion

\*b - offset

\*a - factor

\*CANapeHiddenExtra - sometimes CANape appends extra information; not compliant with MDF specs

-ASAM formula conversion

\*formula - ecuation as string

-polynomial or rational conversion

\*P1 .. P6 - factors

-exponential or logarithmic conversion

\*P1 .. P7 - factors

## **Examples**

## **Attributes**

**address** (int) block address inside mdf file

## **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

## **ChannelDependency Class**

```
class asammdf.mdf3.ChannelDependency(**kargs)
```

CDBLOCK class derived from dict

Currently the ChannelDependency object can only be created using the *file\_stream* and *address* keyword parameters when reading from file

The keys have the following meaning:

- •id Block type identifier, always "CD"
- •block\_len Block size of this block in bytes (entire CDBLOCK)
- •data Dependency type
- •sd\_nr Total number of signals dependencies (m)
- •for each dependency there is a group of three keys:
  - $-dg_{n}$  Pointer to the data group block (DGBLOCK) of signal dependency n
  - $-cg_{n}$  Pointer to the channel group block (DGBLOCK) of signal dependency n
  - $-ch_{n}$  Pointer to the channel block (DGBLOCK) of signal dependency n
- •there can also be optional keys which decribe dimensions for the N-dimensional dependencies:
  - $-\dim_{n}$  Optional: size of dimension n for N-dimensional dependency

## Parameters file\_stream : file handle

mdf file handle

address: int

block address inside mdf file

#### **Attributes**

	address	(int)	block	address	inside	mdf file	
--	---------	-------	-------	---------	--------	----------	--

## **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

## **ChannelExtension Class**

The ChannelExtension object can be created in two modes:

- •using the file\_stream and address keyword parameters when reading from file
- •using any of the following presented keys when creating a new ChannelExtension

The first keys are common for all conversion types, and are followed by conversion specific keys. The keys have the following meaning:

```
•common keys
    -id - Block type identifier, always "CE"
    -block_len - Block size of this block in bytes (entire CEBLOCK)
    -type - Extension type identifier
•specific keys
    -for DIM block
        *module_nr - Number of module
        *module_address - Address
        *description - Description
        *ECU_identification - Identification of ECU
        *reserved0' - reserved
    -for Vector CAN block
        *CAN id - Identifier of CAN message
        *CAN_ch_index - Index of CAN channel
        *message_name - Name of message (string should be terminated by 0)
        *sender_name - Name of sender (string should be terminated by 0)
        *reserved0 - reserved
 Parameters file_stream: file handle
         mdf file handle
     address: int
         block address inside mdf file
```

## **Attributes**

mdf file	inside	address	block	(int)	address
----------	--------	---------	-------	-------	---------

#### **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

## **ChannelGroup Class**

class asammdf.mdf3.ChannelGroup(\*\*kargs)

CGBLOCK class derived from dict

The ChannelGroup object can be created in two modes:

- •using the file\_stream and address keyword parameters when reading from file
- •using any of the following presented keys when creating a new ChannelGroup

The keys have the following meaning:

- •id Block type identifier, always "CG"
- •block\_len Block size of this block in bytes (entire CGBLOCK)
- •next\_cg\_addr Pointer to next channel group block (CGBLOCK) (NIL allowed)
- •first\_ch\_addr Pointer to first channel block (CNBLOCK) (NIL allowed)
- •comment\_addr Pointer to channel group comment text (TXBLOCK) (NIL allowed)
- •record\_id Record ID, i.e. value of the identifier for a record if the DGBLOCK defines a number of record IDs > 0
- •ch\_nr Number of channels (redundant information)
- •samples\_byte\_nr Size of data record in Bytes (without record ID), i.e. size of plain data for a each recorded sample of this channel group
- •cycles\_nr Number of records of this type in the data block i.e. number of samples for this channel group
- •sample\_reduction\_addr only since version 3.3. Pointer to first sample reduction block (SRBLOCK) (NIL allowed) Default value: NIL.

Parameters file\_stream : file handle

mdf file handle

address: int

block address inside mdf file

## **Examples**

#### **Attributes**

**address** (int) block address inside mdf file

## **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

## **DataGroup Class**

```
class asammdf.mdf3.DataGroup(**kargs)
```

DGBLOCK class derived from dict

The DataGroup object can be created in two modes:

- •using the file\_stream and address keyword parameters when reading from file
- •using any of the following presented keys when creating a new DataGroup

The keys have the following meaning:

- •id Block type identifier, always "DG"
- •block\_len Block size of this block in bytes (entire DGBLOCK)
- •next\_dg\_addr Pointer to next data group block (DGBLOCK) (NIL allowed)
- •first\_cg\_addr Pointer to first channel group block (CGBLOCK) (NIL allowed)
- •trigger\_addr Pointer to trigger block (TRBLOCK) (NIL allowed)
- •data\_block\_addr Pointer to the data block (see separate chapter on data storage)
- •cg\_nr Number of channel groups (redundant information)

•record\_id\_nr - Number of record IDs in the data block

•reserved0 - since version 3.2; Reserved

Parameters file\_stream: file handle

mdf file handle

address: int

block address inside mdf file

#### **Attributes**

<b>address</b> (int) block address inside mdf file
--

## **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

## FileIdentificationBlock Class

class asammdf.mdf3.FileIdentificationBlock(\*\*kargs)

IDBLOCK class derived from dict

The TriggerBlock object can be created in two modes:

- •using the *file\_stream* and *address* keyword parameters when reading from file
- •using the classmethod from\_text

The keys have the following meaning:

- •file\_identification file identifier
- •version\_str format identifier
- •program\_identification program identifier
- •byte\_order default byte order
- •float\_format default floating-point format
- •mdf\_version version number of MDF format
- •code\_page code page number
- •reserved0 reserved

```
•reserved1 - reserved
```

•unfinalized\_standard\_flags - Standard Flags for unfinalized MDF

•unfinalized\_custom\_flags - Custom Flags for unfinalized MDF

Parameters file stream: file handle

mdf file handle

version: int

mdf version in case of new file

#### **Attributes**

address	(int) block address inside mdf file; should be 0 always
addi Coo	(int) block address inside mai me, should be o arways

## **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

#### **HeaderBlock Class**

```
class asammdf.mdf3.HeaderBlock(**kargs)
```

HDBLOCK class derived from dict

The TriggerBlock object can be created in two modes:

- •using the *file\_stream* when reading from file
- •using the classmethod from\_text

The keys have the following meaning:

- •id Block type identifier, always "HD"
- •block\_len Block size of this block in bytes (entire HDBLOCK)
- •first\_dg\_addr Pointer to the first data group block (DGBLOCK)
- •comment\_addr Pointer to the measurement file comment text (TXBLOCK) (NIL allowed)
- •program\_addr Pointer to program block (PRBLOCK) (NIL allowed)
- •dg\_nr Number of data groups (redundant information)
- •date Date at which the recording was started in "DD:MM:YYYY" format

```
•time - Time at which the recording was started in "HH:MM:SS" format
```

- •author author name
- •organization organization
- •project project name
- •subject subject

Since version 3.2 the following extra keys were added:

- •abs\_time Time stamp at which recording was started in nanoseconds.
- •tz\_offset UTC time offset in hours (= GMT time zone)
- •time\_quality Time quality class
- •timer\_identification Timer identification (time source),

Parameters file\_stream: file handle

mdf file handle

## **Attributes**

address	(int) block address inside mdf file; should be 64 always
	(int) croth address misrae mar me, should be a cultural s

## **Methods**

clear	
сору	Generic (shallow and deep) copying operations
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

## **ProgramBlock Class**

 ${f class}$  asammdf.mdf3.ProgramBlock(\*\*kargs)

PRBLOCK class derived from dict

The ProgramBlock object can be created in two modes:

- •using the file\_stream and address keyword parameters when reading from file
- •using any of the following presented keys when creating a new ProgramBlock

The keys have the following meaning:

- •id Block type identifier, always "PR"
- •block\_len Block size of this block in bytes (entire PRBLOCK)

•data - Program-specific data

Parameters file\_stream: file handle

mdf file handle

address: int

block address inside mdf file

## **Attributes**

#### **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

## **SampleReduction Class**

class asammdf.mdf3.SampleReduction(\*\*kargs)

SRBLOCK class derived from dict

Currently the SampleReduction object can only be created by using the *file\_stream* and *address* keyword parameters - when reading from file

The keys have the following meaning:

•id - Block type identifier, always "SR"

•block\_len - Block size of this block in bytes (entire SRBLOCK)

•next\_sr\_addr - Pointer to next sample reduction block (SRBLOCK) (NIL allowed)

•data\_block\_addr - Pointer to the data block for this sample reduction

•cycles\_nr - Number of reduced samples in the data block.

•time\_interval - Length of time interval [s] used to calculate the reduced samples.

Parameters file\_stream : file handle

mdf file handle

address: int

block address inside mdf file

## **Attributes**

address	(int	block address inside mdf file
---------	------	-------------------------------

#### **Methods**

clear	
сору	Generic (shallow and deep) copying operations
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

#### **TextBlock Class**

```
class asammdf.mdf3.TextBlock (**kargs)
    TXBLOCK class derived from dict
```

The ProgramBlock object can be created in two modes:

•using the *file\_stream* and *address* keyword parameters - when reading from file

•using the classmethod from\_text

The keys have the following meaning:

```
•id - Block type identifier, always "TX"
```

•block\_len - Block size of this block in bytes (entire TXBLOCK)

•text - Text (new line indicated by CR and LF; end of text indicated by 0)

```
Parameters file_stream : file handle
```

mdf file handle

address: int

block address inside mdf file

text : bytes

bytes for creating a new TextBlock

## **Examples**

```
>>> tx1 = TextBlock.from_text('VehicleSpeed')
>>> tx1.text_str
'VehicleSpeed'
```

```
>>> tx1['text']
b'VehicleSpeed'
```

## **Attributes**

address	(int) block address inside mdf file
text_str	(str) text data as unicode string

## **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
from_text	
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

## **TriggerBlock Class**

```
{\bf class} \; {\tt asammdf.mdf3.TriggerBlock} \; (**kargs)
```

TRBLOCK class derived from dict

The TriggerBlock object can be created in two modes:

- •using the *file\_stream* and *address* keyword parameters when reading from file
- •using the classmethod *from\_text*

The keys have the following meaning:

- •id Block type identifier, always "TX"
- •block\_len Block size of this block in bytes (entire TRBLOCK)
- •text\_addr Pointer to trigger comment text (TXBLOCK) (NIL allowed)
- •trigger\_events\_nr Number of trigger events n (0 allowed)
- •trigger\_{n}\_time Trigger time [s] of trigger event n
- •trigger\_ $\{n\}$ \_pretime Pre trigger time [s] of trigger event n
- •trigger\_ $\{n\}$ \_posttime Post trigger time [s] of trigger event n

Parameters file\_stream : file handle

mdf file handle

address: int

block address inside mdf file

## **Attributes**

block address inside mdf file	address
-------------------------------	---------

## **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

#### MDF4

asammdf tries to emulate the mdf structure using Python builtin data types.

The header attibute is an OrderedDict that holds the file metadata.

The *groups* attribute is a dictionary list with the following keys:

- data\_group : DataGroup object
- channel\_group : ChannelGroup object
- channels: list of Channel objects with the same order as found in the mdf file
- channel\_conversions : list of ChannelConversion objects in 1-to-1 relation with the channel list
- channel\_sources : list of SourceInformation objects in 1-to-1 relation with the channels list
- data\_block : DataBlock object
- texts: dictionay containing TextBlock objects used throughout the mdf
  - channels: list of dictionaries that contain TextBlock objects ralated to each channel
    - \* name\_addr : channel name
    - \* comment\_addr : channel comment
  - channel group: list of dictionaries that contain TextBlock objects ralated to each channel group
    - \* acq\_name\_addr : channel group acquisition comment
    - \* comment\_addr : channel group comment
  - conversion\_tab: list of dictionaries that contain TextBlock objects related to TABX and RTABX channel conversions
    - \* text\_{n}: n-th text of the VTABR conversion

- \* default addr : default text
- conversions : list of dictionaries that containt TextBlock obejets related to channel conversions

\* name\_addr : converions name

\* unit\_addr : channel unit\_addr

\* comment\_addr : converison comment

\* formula\_addr : formula text; only valid for algebraic conversions

- sources : list of dictionaries that containt TextBlock obejcts related to channel sources

\* name\_addr : source name

\* path\_addr : source path\_addr

\* comment\_addr : source comment

The *file\_history* attribute is a list of (FileHistory, TextBlock) pairs .

The *channel\_db* attibute is a dictionary that holds the *(data group index, channel index)* pair for all signals. This is used to speed up the *get\_signal\_by\_name* method.

The *master\_db* attibute is a dictionary that holds the *channel index* of the master channel for all data groups. This is used to speed up the *get\_signal\_by\_name* method.

#### API

class asammdf.mdf4.MDF4 (name=None, load\_measured\_data=True, version='4.00')

If the name exist it will be loaded otherwise an empty file will be created that can be later saved to disk

Parameters name: string

mdf file name

load\_measured\_data : bool

load data option; default True

- if True the data group binary data block will be loaded in RAM
- if False the channel data is read from disk on request

version: string

mdf file version ('4.00', '4.10', '4.11'); default '4.00'

#### **Attributes**

name	(string) mdf file name	
groups	(list) list of data groups	
header	(HeaderBlock) mdf file header	
file_history	(list) list of (FileHistory, TextBlock) pairs	
comment	(TextBlock) mdf file comment	
identification	(FileIdentificationBlock) mdf file start block	
load_measured_d	load_measured_dat@bool) load measured data option	
version	(int) mdf version	
channels_db	(dict) used for fast channel access by name; for each name key the value is a (group	
	index, channel index) tuple	
masters_db	(dict) used for fast master channel access; for each group index key the value is the	
	master channel index	

# **Methods**

append	
attach	
extract_attachment	
get	
nfo	
remove	
save	

```
append (signals, source_info='Python', common_timebase=False)
Appends a new data group.

Parameters signals: list
    list on Signal objects
acquisition_info: str
    acquisition information; default 'Python'
common_timebase: bool
flag to hint that the signals have the same timebase
```

#### **Examples**

```
>>> # case 1 conversion type None
>>> s1 = np.array([1, 2, 3, 4, 5])
>>> s2 = np.array([-1, -2, -3, -4, -5])
>>> s3 = np.array([0.1, 0.04, 0.09, 0.16, 0.25])
>>> t = np.array([0.001, 0.002, 0.003, 0.004, 0.005])
>>> names = ['Positive', 'Negative', 'Float']
>>> units = ['+', '-', '.f']
>>> info = {}
>>> s1 = Signal(samples=s1, timstamps=t, unit='+', name='Positive')
>>> s2 = Signal(samples=s2, timstamps=t, unit='-', name='Negative')
>>> s3 = Signal(samples=s3, timstamps=t, unit='flts', name='Floats')
```

```
>>> mdf = MDF4('new.mf4')
>>> mdf.append([s1, s2, s3], 'created by asammdf v1.1.0')
>>> # case 2: VTAB conversions from channels inside another file
>>> mdf1 = MDF4('in.mf4')
>>> ch1 = mdf1.get("Channel1_VTAB")
>>> ch2 = mdf1.get("Channel2_VTABR")
>>> sigs = [ch1, ch2]
>>> mdf2 = MDF4('out.mf4')
>>> mdf2.append(sigs, 'created by asammdf v1.1.0')
```

attach embedded attachment as application/octet-stream

# Parameters data: bytes

data to be attached

 $\mathbf{file\_name}: str$ 

comment: str

string file name

attachment comment

compression: bool

use compression for embedded attachment data

mime: str

mime type string

#### extract\_attachment (index)

extract attachemnt *index* data. If it is an embedded attachment, then this method creates the new file according to the attachemnt file name information

Parameters index : int
attachment index

Returns data : bytes | str

attachment data

**get** (name=None, group=None, index=None, raster=None, samples\_only=False) Gets channel samples. Channel can be specified in two ways:

•using the first positional argument *name* 

- -if there are multiple occurances for this channel then the *group* argument can be used to select a specific group.
- -if there are multiple occurances for this channel and the *group* argument is None then a warning is issued

•using the group number (keyword argument *group*) and the channel number (keyword argument *in-dex*). Use *info* method for group and channel numbers

If the raster keyword argument is not None the output is interpolated accordingly

Parameters name: string name of channel

```
group: int
                0-based group index
             index: int
                0-based channel index
              raster: float
                time raster in seconds
             samples_only: bool
                if True return only the channel samples as numpy array; if False return a Signal object
         Returns res: (numpy.array | Signal)
                returns Signal if samples_only*=*False (default option), otherwise returns numpy.array
         Raises MdfError:
             * if the channel name is not found
             * if the group index is out of range
             * if the channel index is out of range
info()
     get MDF information as a dict
```

# **Examples**

```
>>> mdf = MDF4('test.mdf')
>>> mdf.info()
```

remove (group=None, name=None)

Remove data group. Use *group* or *name* keyword arguments to identify the group's index. *group* has priority

Parameters name: string

name of the channel inside the data group to be removed

group: int

data group index to be removed

# **Examples**

```
>>> mdf = MDF4('test.mdf')
>>> mdf.remove(group=3)
>>> mdf.remove(name='VehicleSpeed')
```

save (dst=None)

Save MDF to dst. If dst is None the original file is overwritten

#### MDF version 4 blocks

The following classes implement different MDF version3 blocks.

# **AttachmentBlock Class**

class asammdf.mdf4.AttachmentBlock(\*\*kargs)

ATBLOCK class

When adding new attachments only embedded attachemnts are allowed, with keyword argument *data* of type bytes

#### **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
extract	
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

# **Channel Class**

class asammdf.mdf4.Channel(\*\*kargs)

CNBLOCK class

# **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

# **ChannelConversion Class**

 $\begin{array}{c} \textbf{class} \; \texttt{asammdf.mdf4.ChannelConversion} \; (**kargs) \\ \text{CCBLOCK class} \end{array}$ 

# **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

# **ChannelGroup Class**

 ${\bf class} \; {\tt asammdf.mdf4.ChannelGroup} \; (\; {**kargs}) \\$ 

CGBLOCK class

# **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

# **DataGroup Class**

 ${\bf class} \; {\tt asammdf.mdf4.DataGroup} \; (**kargs)$ 

DGBLOCK class

# **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
	Continued on next page

Table 6.20 – continued from previous page

pop	
popitem	
setdefault	
update	
values	

# **DataList Class**

```
class asammdf.mdf4.DataList (**kargs)
    DLBLOCK class
```

# **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

# **DataBlock Class**

Parameters address: int

DTBLOCK address inside the file

file\_stream : int
 file handle

# **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
	Continued on next page

# Table 6.22 – continued from previous page

setdefault			
update			
values			

# FileIdentificationBlock Class

 ${\bf class} \ {\tt asammdf.mdf4.FileIdentificationBlock} \ (**kargs) \\ {\bf IDBLOCK} \ {\bf class} \\$ 

#### **Methods**

clear	
сору	Generic (shallow and deep) copying operations
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

# **HeaderBlock Class**

#### **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

# **SourceInformation Class**

# **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

# **FileHistory Class**

 ${\bf class} \; {\tt asammdf.mdf4.FileHistory} \; (**kargs)$ 

FHBLOCK class

# **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
fromkeys	
get	
items	
keys	
pop	
popitem	
setdefault	
update	
values	

# **TextBlock Class**

# **Methods**

clear	
сору	Generic (shallow and deep) copying operations.
from_text	
fromkeys	
get	
items	
	Continued on next page

Table 6.27 – continued from previous page

keys	
pop	
popitem	
setdefault	
update	
values	

#### classmethod from\_text (text, meta=False)

Create a TextBlock from a str or bytes

Parameters text: str | bytes

input text

meta: bool

enable meta text block

# **Examples**

```
>>> t = TextBlock.from_text(b'speed')
>>> t['id']
b'##TX'
>>> t.text_str
speed
>>> t = TextBlock.from_text('mass', meta=True)
>>> t['id']
b'##MD'
```

# Notes about *load\_measured\_data* argument

By default when the *MDF* object is created the raw channel data is loaded into RAM. This will give you the best performance from *asammdf*.

However if you reach the physical memmory limit *asammdf* gives you the option use the *load\_measured\_data* flag. In this case the raw channel data is not read.

# **MDF** defaults

# Advantages

· best performance

#### Disadvantages

• higher RAM usage, there is the chance the file will exceed available RAM

#### Use case

• when data fits inside the system RAM

# MDF with load\_measured\_data

#### Advantages

- · lowest RAM usage
- faster than compression

#### Disadvantages

• slow performance for getting channel data

#### Use case

· when default data exceeds available RAM

**Note:** See benchmarks for the effects of using the flag

# **Signal**

The Signal represents a signal described by it's samples and timestamps. It can do aritmethic operations agains other Signal or numeric type. The operations are computed in respect to the timestamps (time correct). The integer signals are not interpolated, instead the last value relative to the current timestamp is used. *samples*, *timstamps* and *name* are mandatory arguments.

```
Parameters samples: numpy.array
signal samples
timestamps: numpy.array
signal timestamps
unit: str
signal unit
name: str
signal name
conversion: dict
dict describing the channel conversion, default None
comment: str
signal comment, default ''
```

#### **Methods**

astype			
cut			
interp			
plot			

6.2. Signal 43

astype (np\_type)

```
returns new Signal with samples of dtype np_type

cut (start, stop)

Cuts the signal according to the start and stop values, by using the insertion indexes in the signal's time axis.

Parameters start: float

start timestamp for cutting

stop: float

stop timestamp for cutting

Returns outsig: Signal

new Signal cut from the original
```

# **Examples**

```
>>> new_sig = old_sig.cut(1.0, 10.5)
>>> new_sig.timestamps[0], new_sig.timestamps[-1]
0.98, 10.48

interp(new_timestamps)
   returns a new Signal interpolated using the new_timestamps

plot()
   plot Signal samples
```

# **Examples**

# Working with MDF

```
from asammdf import MDF, Signal
import numpy as np
# create 3 Signal objects
timestamps = np.array([0.1, 0.2, 0.3, 0.4, 0.5], dtype=np.float32)
# unit8
s_uint8 = Signal(samples=np.array([0, 1, 2, 3, 4], dtype=np.uint8),
                 timestamps=timestamps,
                 name='Uint8_Signal',
                 unit='u1')
# int32
s_{int32} = Signal(samples=np.array([-20, -10, 0, 10, 20], dtype=np.int32),
                 timestamps=timestamps,
                 name='Int32_Signal',
                 unit='i4')
# float64
s_{10at64} = Signal(samples=np.array([-20, -10, 0, 10, 20], dtype=np.int32),
```

```
timestamps=timestamps,
                   name='Float64_Signal',
                   unit='f8')
# create empty MDf version 4.00 file
mdf4 = MDF(version='4.00')
# append the 3 signals to the new file
signals = [s_uint8, s_int32, s_float64]
mdf4.append(signals, 'Created by Python')
# save new file
mdf4.save('my_new_file.mf4')
# convert new file to mdf version 3.10 with compression of raw channel data
mdf3 = mdf4.convert(to='3.10', compression=True)
print (mdf3.version)
# prints >>> 3.10
# get the float signal
sig = mdf3.get('Float64_Signal')
print(sig)
                                                  s=[-20 \ -10 \ 0 \ 10 \ 20] \ t=[\ 0.1 \ ]
# prints >>> Signal { name="Float64_Signal":
                   0.30000001 0.40000001 0.5
                                                   ] unit="f8"
      0.2
→conversion=None }
```

# **Working with Signal**

```
from asammdf import Signal
import numpy as np
# create 3 Signal objects with different time stamps
# unit8 with 100ms time raster
timestamps = np.array([0.1 * t for t in range(5)], dtype=np.float32)
s_uint8 = Signal(samples=np.array([t for t in range(5)], dtype=np.uint8),
                 timestamps=timestamps,
                 name='Uint8_Signal',
                 unit='u1')
# int32 with 50ms time raster
timestamps = np.array([0.05 * t for t in range(10)], dtype=np.float32)
s_int32 = Signal(samples=np.array(list(range(-500, 500, 100)), dtype=np.int32),
                 timestamps=timestamps,
                 name='Int32_Signal',
                 unit='i4')
# float64 with 300ms time raster
timestamps = np.array([0.3 * t for t in range(3)], dtype=np.float32)
s_float64 = Signal(samples=np.array(list(range(2000, -1000, -1000)), dtype=np.int32),
                   timestamps=timestamps,
                   name='Float64_Signal',
                   unit='f8')
prod = s_float64 * s_uint8
```

6.3. Examples 45

```
prod.name = 'Uint8_Signal * Float64_Signal'
prod.unit = '*'
prod.plot()

pow2 = s_uint8 ** 2
pow2.name = 'Uint8_Signal ^ 2'
pow2.unit = 'u1^2'
pow2.plot()

allsum = s_uint8 + s_int32 + s_float64
allsum.name = 'Uint8_Signal + Int32_Signal + Float64_Signal'
allsum.unit = '+'
allsum.plot()

# inplace operations
pow2 *= -1
pow2.name = '- Uint8_Signal ^ 2'
pow2.plot()
```

# **Benchmarks**

asammdf relies heavily on dict objects. Starting with Python 3.6 the dict objects are more compact and ordered (implementation detail); asammdf uses takes advantage of those changes so for best performance it is advised to use Python >= 3.6.

# Intro

The benchmarks were done using two test files (for mdf version 3 and 4) of around 170MB. The files contain 183 data groups and a total of 36424 channels.

asamdf 2.3.2 was compared against mdfreader 0.2.5. mdfreader seems to be the most used Python package to handle MDF files, and it also supports both version 3 and 4 of the standard.

The three benchmark cathegories are file open, file save and extracting the data for all channels inside the file(36424 calls). For each cathegory two aspect were noted: elapsed time and peak RAM usage.

# **Dependencies**

You will need the following packages to be able to run the benchmark script

- psutil
- mdfreader

# x64 Python results

The test environment used for 64 bit tests had:

- 3.6.2 (v3.6.2:5fd33b5, Jul 8 2017, 04:57:36) [MSC v.1900 64 bit (AMD64)]
- Windows-10-10.0.14393-SP0

- Intel64 Family 6 Model 94 Stepping 3, GenuineIntel
- 16GB installed RAM

Notations used in the results

- nodata = MDF object created with load\_measured\_data=False (raw channel data not loaded into RAM)
- compression = MDF object created with compression=blosc
- compression bcolz 6 = MDF object created with compression=6
- noDataLoading = MDF object read with noDataLoading=True

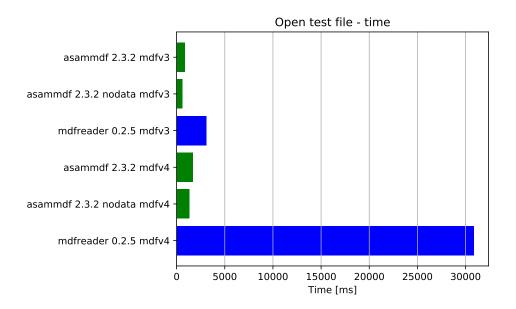
Files used for benchmark: \* 183 groups \* 36424 channels

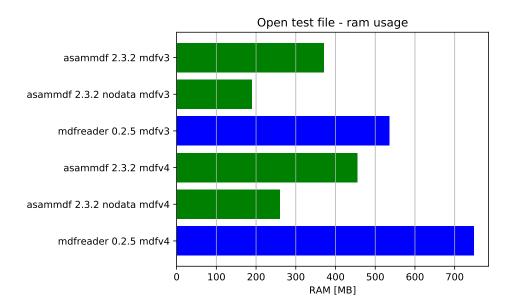
Open file	Time [ms]	RAM [MB]
asammdf 2.3.2 mdfv3	831	371
asammdf 2.3.2 nodata mdfv3	609	190
mdfreader 0.2.5 mdfv3	3083	536
asammdf 2.3.2 mdfv4	1710	455
asammdf 2.3.2 nodata mdfv4	1349	260
mdfreader 0.2.5 mdfv4	30847	748

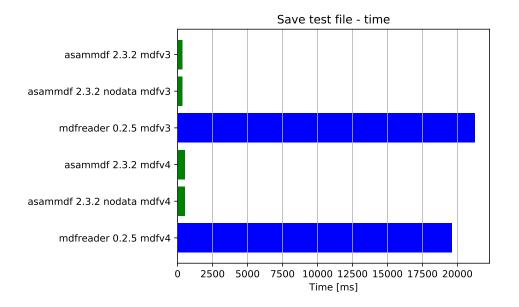
Save file	Time [ms]	RAM [MB]
asammdf 2.3.2 mdfv3	348	371
asammdf 2.3.2 nodata mdfv3	343	197
mdfreader 0.2.5 mdfv3	21244	1997
asammdf 2.3.2 mdfv4	530	462
asammdf 2.3.2 nodata mdfv4	522	272
mdfreader 0.2.5 mdfv4	19594	2795

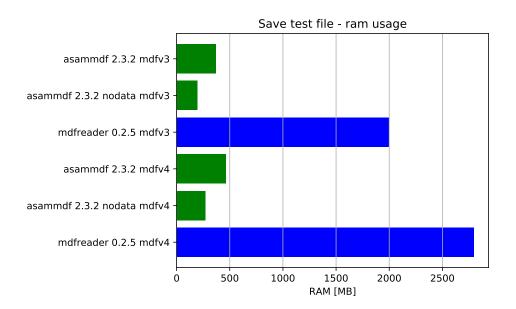
Get all channels (36424 calls)	Time [ms]	RAM [MB]
asammdf 2.3.2 mdfv3	681	383
asammdf 2.3.2 nodata mdfv3	9175	209
mdfreader 0.2.5 mdfv3	29	537
asammdf 2.3.2 mdfv4	599	464
asammdf 2.3.2 nodata mdfv4	12191	273
mdfreader 0.2.5 mdfv4	38	748

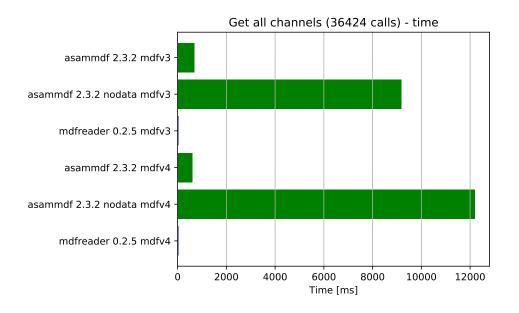
# **Graphical results**

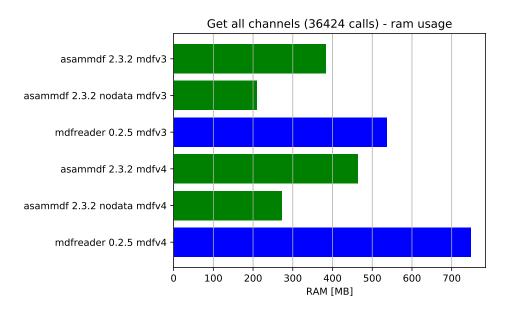












# x86 Python results

The test environment used for 32 bit tests had:

- Python 3.6.1 (v3.6.1:69c0db5, Mar 21 2017, 17:54:52) [MSC v.1900 32 bit (Intel)]
- Windows-7-6.1.7601-SP1
- Intel64 Family 6 Model 94 Stepping 3, GenuineIntel (i7-6820Q)
- 16GB installed RAM

The notations used in the results have the following meaning:

- nodata = MDF object created with load\_measured\_data=False (raw channel data no loaded into RAM)
- compression = MDF object created with compression=True (raw channel data loaded into RAM and compressed)
- noconvert = MDF object created with convertAfterRead=False

# Raw data

- 3.6.1 (v3.6.1:69c0db5, Mar 21 2017, 17:54:52) [MSC v.1900 32 bit (Intel)]
- Windows-10-10.0.14393-SP0
- Intel64 Family 6 Model 94 Stepping 3, GenuineIntel
- 16GB installed RAM

Notations used in the results

- nodata = MDF object created with load\_measured\_data=False (raw channel data not loaded into RAM)
- compression = MDF object created with compression=True/blosc
- compression boolz 6 = MDF object created with compression=6
- noDataLoading = MDF object read with noDataLoading=True

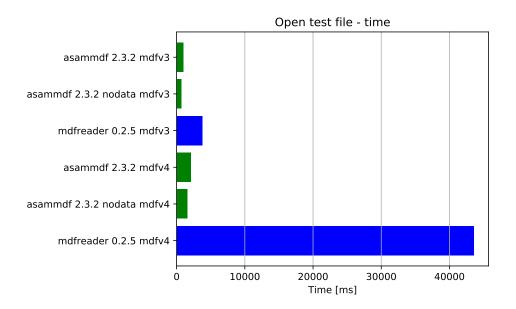
Files used for benchmark: \* 183 groups \* 36424 channels

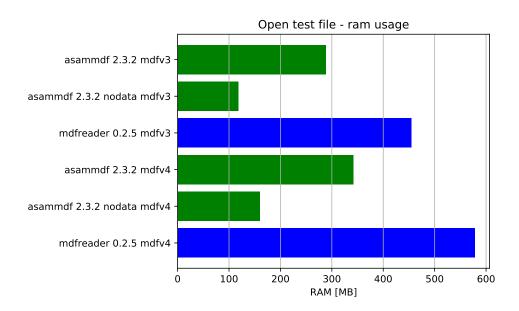
Open file	Time [ms]	RAM [MB]
asammdf 2.3.2 mdfv3	980	288
asammdf 2.3.2 nodata mdfv3	670	118
mdfreader 0.2.5 mdfv3	3776	455
asammdf 2.3.2 mdfv4	2071	342
asammdf 2.3.2 nodata mdfv4	1610	160
mdfreader 0.2.5 mdfv4	43559	578

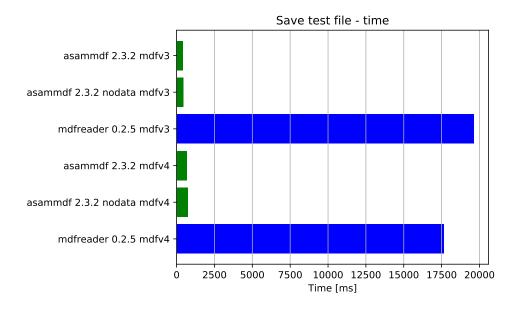
Save file	Time [ms]	RAM [MB]
asammdf 2.3.2 mdfv3	406	291
asammdf 2.3.2 nodata mdfv3	432	125
mdfreader 0.2.5 mdfv3	19623	1224
asammdf 2.3.2 mdfv4	691	351
asammdf 2.3.2 nodata mdfv4	734	169
mdfreader 0.2.5 mdfv4	17657	1687

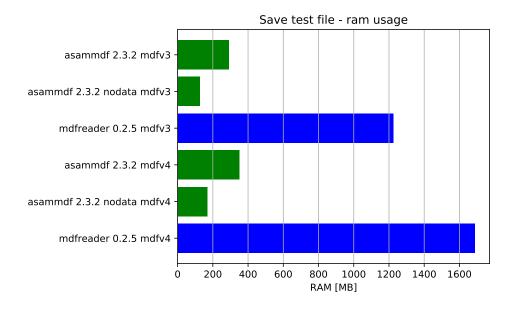
Get all channels (36424 calls)	Time [ms]	RAM [MB]
asammdf 2.3.2 mdfv3	963	298
asammdf 2.3.2 nodata mdfv3	19059	132
mdfreader 0.2.5 mdfv3	34	455
asammdf 2.3.2 mdfv4	868	349
asammdf 2.3.2 nodata mdfv4	20434	171
mdfreader 0.2.5 mdfv4	54	578

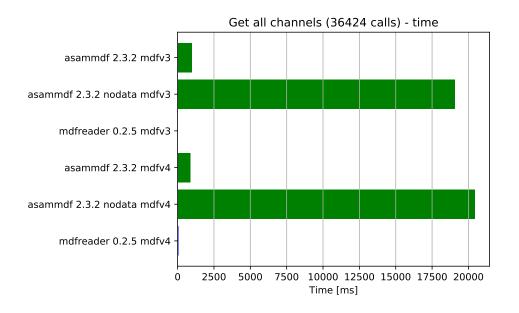
# **Graphical results**

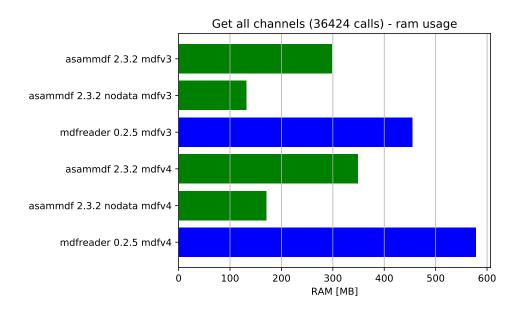












# CHAPTER 8

# Indices and tables

- genindex
- modindex
- search

# Index

A	Н
add_trigger() (asammdf.mdf3.MDF3 method), 15 append() (asammdf.mdf3.MDF3 method), 16	HeaderBlock (class in asammdf.mdf3), 27 HeaderBlock (class in asammdf.mdf4), 40
append() (asammdf.mdf4.MDF4 method), 34 astype() (asammdf.signal.Signal method), 44	1
attach() (asammdf.mdf4.MDF4 method), 35 AttachmentBlock (class in asammdf.mdf4), 37	info() (asammdf.mdf3.MDF3 method), 17 info() (asammdf.mdf4.MDF4 method), 36
C	interp() (asammdf.signal.Signal method), 44 iter_get_triggers() (asammdf.mdf3.MDF3 method), 17
Channel (class in asammdf.mdf3), 18 Channel (class in asammdf.mdf4), 37	iter_to_pandas() (asammdf.mdf.MDF method), 14
ChannelConversion (class in asammdf.mdf3), 20	M
ChannelConversion (class in asammdf.mdf4), 37 ChannelDependency (class in asammdf.mdf3), 22 ChannelExtension (class in asammdf.mdf3), 23	MDF (class in asammdf.mdf), 13 MDF3 (class in asammdf.mdf3), 15
ChannelGroup (class in asammdf.mdf3), 24	MDF4 (class in asammdf.mdf4), 33
ChannelGroup (class in asammdf.mdf4), 38	P
convert() (asammdf.mdf.MDF method), 13 cut() (asammdf.signal.Signal method), 44	plot() (asammdf.signal.Signal method), 44 ProgramBlock (class in asammdf.mdf3), 28
D	R
DataBlock (class in asammdf.mdf4), 39 DataGroup (class in asammdf.mdf3), 25 DataGroup (class in asammdf.mdf4), 38	remove() (asammdf.mdf3.MDF3 method), 17 remove() (asammdf.mdf4.MDF4 method), 36
DataList (class in asammdf.mdf4), 39	S
E	SampleReduction (class in asammdf.mdf3), 29
extract_attachment() (asammdf.mdf4.MDF4 method), 35	save() (asammdf.mdf3.MDF3 method), 18 save() (asammdf.mdf4.MDF4 method), 36
F	Signal (class in asammdf.signal), 43
FileHistory (class in asammdf.mdf4), 41	SourceInformation (class in asammdf.mdf4), 40
FileIdentificationBlock (class in asammdf.mdf3), 26	T
FileIdentificationBlock (class in asammdf.mdf4), 40	TextBlock (class in asammdf.mdf3), 30
filter() (asammdf.mdf.MDF method), 14 from_text() (asammdf.mdf4.TextBlock class method), 42	TextBlock (class in asammdf.mdf4), 41 TriggerBlock (class in asammdf.mdf3), 31
G	
get() (asammdf.mdf3.MDF3 method), 16	
get() (asammdf.mdf4.MDF4 method), 35	