Array Methods HDF5 File Definitions

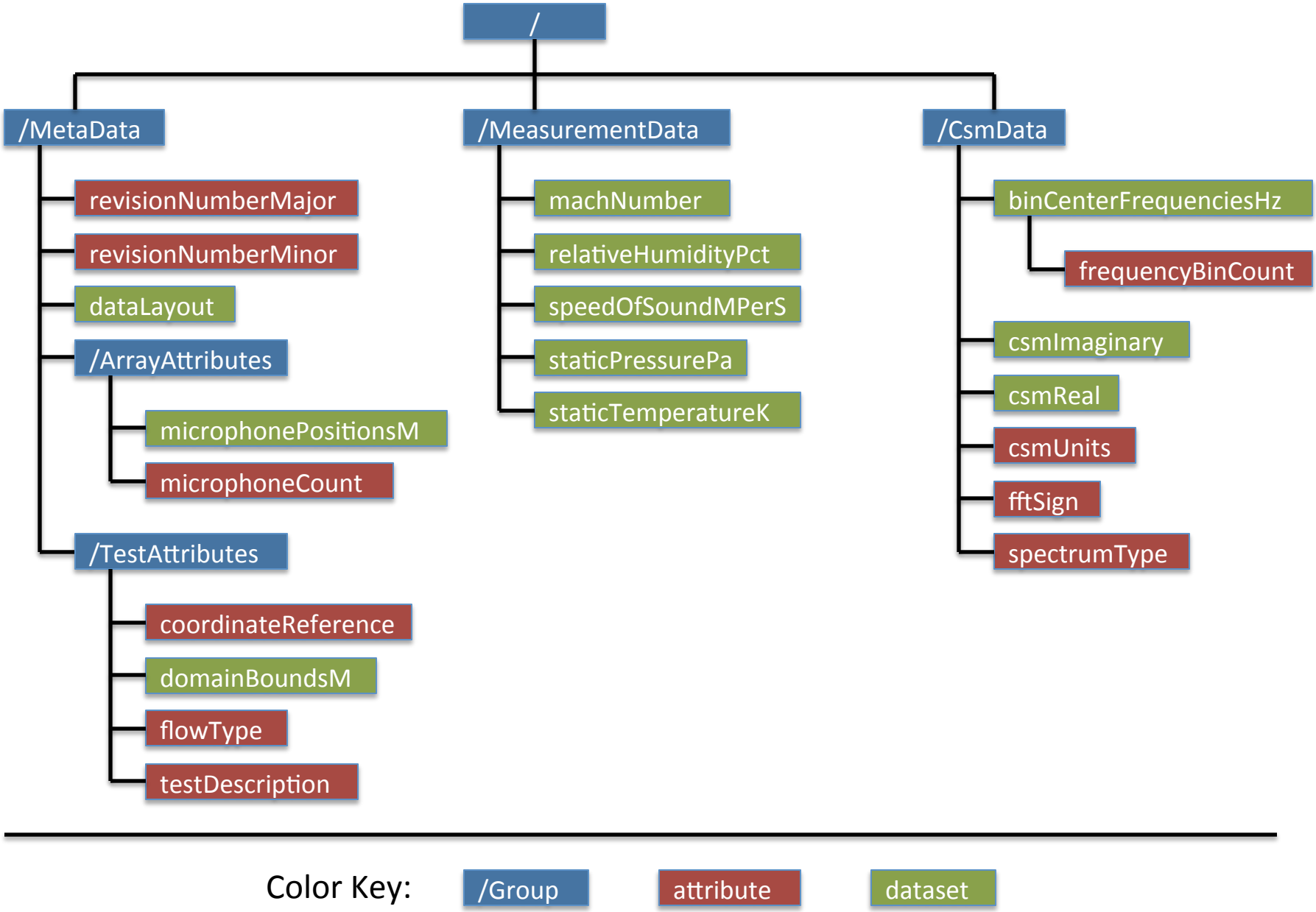
Revision 2.4 Release

September 1, 2016

CSM Format: Purpose and Files

- The CSM file standard exists to streamline distribution and use of data for the array methods benchmarking effort
- HDF5 file standard selected to minimize impact of choice of analysis platforms
- The CSM file standard will define the near-minimum data required to conduct analysis in the frequency domain
- A secondary HDF5 file can be generated for a given case to include optional information, e.g. an example solution with one type of steering vector choice on a given grid, or possibly a visualization mesh. The included example optional CSM is one way in which this can be formatted. This could also include additional information for varying physical properties, e.g. in-stream vs. out-of-stream speed of sound. Associated documentation will be required to describe the contents.
- The file with the required data for a case named <caseID> will be named <caseID>CsmEss.h5, indicating it is "essential." A file with optional data will be named <caseID>CsmOpt.h5, indicating it is "optional"
- It has been decided that including additional details in the CSM files regarding shear layers, reflective surfaces, etc. is too cumbersome to standardize given the variety of conditions, as well as individual user needs. These must instead be described in attached documentation, with sufficient detail for the user to model. For example, if a facility has a single rectangular, planar shear layer, what are the coordinates of its corner vertices in the selected coordinate system?
- Case documentation will be named <caseID>Info.<pdf,doc,docx,txt…>
- Notes:
  - Revision 2.3 incorporates the dataLayout dataset, which should remove any ambiguity in loading an HDF5 file.

Array Methods HDF5 Essential CSM File Layout



Array Methods HDF5 Essential CSM File Definitions

- /MetaData
  - revisionNumberMajor
    - major revision number of standard to which the file adheres
    - format: 32-bit signed integer
  - revisionNumberMinor
    - minor revision number of standard to which the file adheres
    - format: 32-bit signed integer
  - dataLayout
    - example three-dimensional data structure for validating data read orientation
    - size 2 x 3 x 4 (rows x columns x pages)
    - holds increasing integers in column-major order:
      page 1 – [1 3 5 / 2 4 6] page 2 – [7 9 11 / 8 10 12] page 3 – [13 15 17 / 14 16 18] page 4 – [19 21 23 / 20 22 24]
    - format: 32-bit signed integer
  - /ArrayAttributes
    - microphonePositionsM
      - array of microphone coordinates
      - size: microphoneCount x 3
      - units: meters
      - format: floating point
      - note: each row of the array corresponds to a single microphone's coordinate triple as <x,y,z>
    - microphoneCount
      - number of microphones in the microphone array
      - format: 32-bit signed integer

Array Methods HDF5 Essential CSM File Definitions

- /MetaData
  - /TestAttributes
    - coordinateReference
      - string describing coordinate reference frame and/or origin
      - examples: `array center', `test section entrance'
    - domainBoundsM
      - bounds of the volume containing the source region of interest
      - size: 2 x 3; rows are domain minima and maxima; columns are x-, y- and z-bounds
      - units: meters
      - format: floating point
    - flowType
      - string describing flow field of simulation/measurement
      - examples: `no flow', `uniform flow', `open jet'
    - testDescription
      - string outlining details of simulation/measurement
      - examples: `simulated line source in open jet', `Date; Place; Test point no.; Condition'

Array Methods HDF5 Essential CSM File Definitions

- /MeasurementData
  - machNumber
    - Mach number of flow field
    - size: 1 x 3, corresponding to x-, y- and z-components
    - format: floating point
  - relativeHumidityPct
    - relative humidity of simulation/measurement
    - units: percent
    - format: floating point
  - speedOfSoundMPerS
    - speed of sound of simulation/measurement
    - units: m/s
    - format: floating point
  - staticPressurePa
    - static pressure of simulation/measurement
    - units: Pa
    - format: floating point
  - staticTemperatureK
    - Static temperature of simulation/measurement
    - units: Kelvin
    - format: floating point

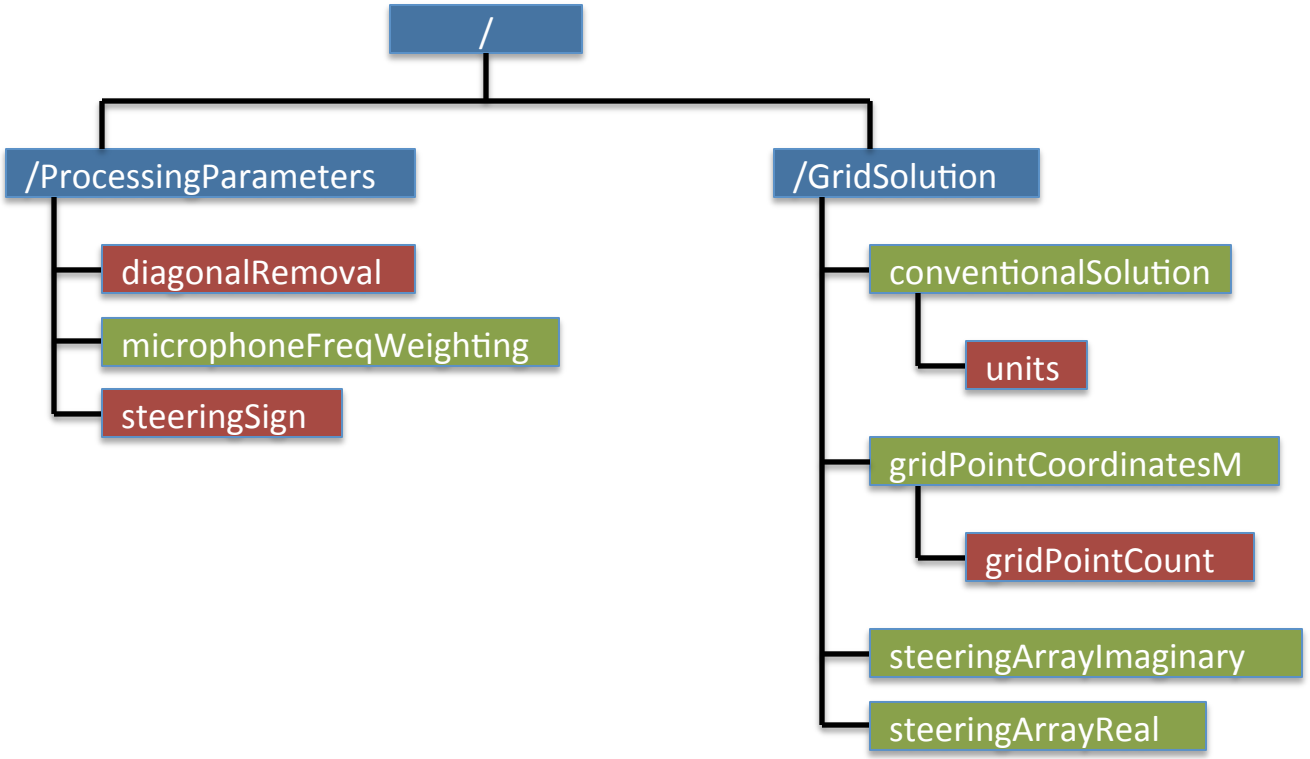
Array Methods HDF5 Essential CSM File Definitions

- /CsmData
  - binCenterFrequenciesHz
    - array of center frequencies of CSM frequency bins
    - size: 1 x frequencyBinCount
    - units: Hz
    - format: floating point
    - note: assume bin boundaries are halfway between elements of the array for narrowband/ power spectral density; follow base-10 formulas in ANSI S1.11-2004 (or ISO equivalent) for 1/ n^{th}-octave centers and edges
    - frequencyBinCount
      - number of frequency bins in spectral processing
      - format: 32-bit signed integer
  - csmImaginary
    - imaginary component of the CSM data array
    - size: microphoneCount x microphoneCount x frequencyBinCount
    - units: Pa^2 for narrowband/fractional octave, Pa^2/Hz for power spectral density
    - format: floating point
    - chunk size: : microphoneCount x microphoneCount x 1
    - note: each chunk should be an antisymmetric matrix with a zero diagonal
    - note: sign convention should correspond to construction via the outer product of the column vector of microphone Fourier transform terms for the given frequency bin and the hermitian transpose of that column vector

Array Methods HDF5 Essential CSM File Definitions

- /CsmData
  - csmReal
    - real component of the CSM data array
    - size: microphoneCount x microphoneCount x frequencyBinCount
    - units: Pa^2 for narrowband/fractional octave, Pa^2/Hz for power spectral density
    - format: floating point
    - chunk size: : microphoneCount x microphoneCount x 1
    - note: each chunk should be a symmetric matrix
    - note: sign convention should correspond to construction via the outer product of the column vector of microphone Fourier transform terms for the given frequency bin and the hermitian transpose of that column vector
  - csmUnits
    - string stating units of CSM data array
    - entries: `Pa^2' for spectrumType `narrowband' and `octave-n', `Pa^2/Hz' for spectrumType `psd'
  - fftSign
    - sign of the exponent used to Fourier transform the data ($i\omega t$ vs. $-i\omega t$)
    - format: 32-bit signed integer
    - value: +/- 1
  - spectrumType
    - String describing type of spectrum in CSM data array
    - Examples: `narrowband', `psd' for power spectral density, and `octave-n' for $1/n^{th}$-octave

Array Methods HDF5 Example Optional CSM File



Color Key: /Group attribute dataset

Array Methods HDF5 Example Optional CSM File

- /ProcessingParameters
  - diagonalRemoval
    - string stating if diagonal removal was applied in computing conventionalSolution
    - options: `true', `false'
  - microphoneFreqWeighting
    - microphone frequency-dependent weighting coefficients used in computing conventionalSolution
    - size: microphoneCount x frequencyBinCount
    - format: floating point
    - chunk size: microphoneCount x 1
    - note: for no weighting, set all elements to unity; to disable microphones, e.g. use a subarray, set the disabled microphone weights to zero
  - steeringSign
    - sign of the exponent used to construct the steering vectors (*ikr* vs. *–ikr*)
    - format: 32-bit signed integer, value +/- 1

Array Methods HDF5 Example Optional CSM File

- /GridSolution
  - conventionalSolution
    - conventional beamforming solution
    - size: gridPointCount x frequencyBinCount
    - format: floating point
    - chunk size: gridPointCount x 1
    - note: conventional beamforming solution using the included matrix of steering vectors, microphone weighting, and diagonal removal setting
    - units
      - string giving units for conventionalSolution
      - examples: `Pa^2' or `dB ref 20 microPa'
  - gridPointCoordinatesM
    - array of grid point coordinates for conventional beamforming solution
    - size: gridPointCount x 3
    - units: meters
    - format: floating point
    - note: each row of the array corresponds to a grid point's coordinate triple as <x,y,z>
    - gridPointCount
      - number of scan grid points for optional steering vectors and solution
      - format: 32-bit signed integer

Array Methods HDF5 Example Optional CSM File

- /GridSolution
  - steeringArrayImaginary
    - imaginary component of array of steering vectors relating each microphone to each grid point for every frequency
    - size: microphoneCount x gridPointCount x frequencyBinCount
    - format: floating point
    - chunk size: microphoneCount x gridPointCount x 1
    - note: each column of a chunk relates all microphones to a single grid point
  - steeringArrayReal
    - real component of array of steering vectors relating each microphone to each grid point for every frequency
    - size: microphoneCount x gridPointCount x frequencyBinCount
    - format: floating point
    - chunk size: microphoneCount x gridPointCount x 1
    - note: each column of a chunk relates all microphones to a single grid point

Time Series Format: Purpose and Files

- The time series file standard will define the storage of the raw pressure time histories, as well as the minimum information required to construct the CSM of the data. The intent of the time series file is to augment the supplied CSM and documentation of a data set for further advanced analysis, not to replace it.

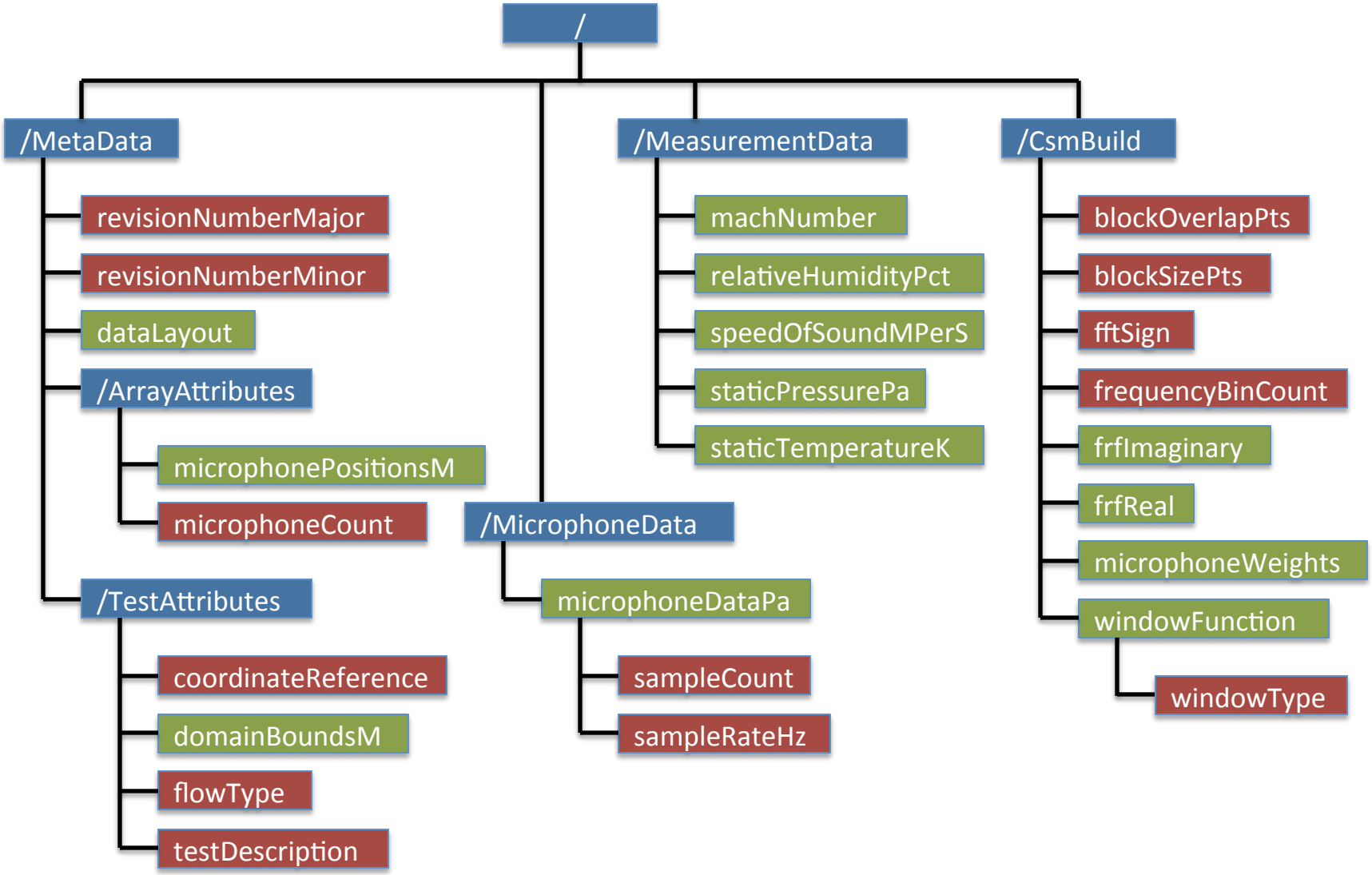
- /MetaData and /MeasurementData are duplicated from the CSM file for convenience.

- A secondary HDF5 file can be generated for a given case to include additional information, e.g. additional sensor channels for speaker input, tachometer signals, IRIG, etc. The included example file is one way in which this could be formatted. Associated documentation is necessary to describe the contents.

- The file with the required time series data for a case named <caseID> will be named <caseID>TimeSeries.h5. A file with the optional data will be named <caseID> TimeSeriesOpt.h5, indicating it has optional additional information

- Notes:
  - Revision 2.3 incorporates the dataLayout dataset, which should remove any ambiguity in loading an HDF5 file.
  - Revision 2.4 transposes the microphoneDataPa array layout. User experience showed that, based on typical data usage, the previous array layout was a poor choice from a read/write performance perspective.

Array Methods HDF5 Time Series File Layout



Color Key: /Group attribute dataset

Array Methods HDF5 Time Series File Definitions

- /MetaData
  - revisionNumberMajor
    - major revision number of standard to which the file adheres
    - format: 32-bit signed integer
  - revisionNumberMinor
    - minor revision number of standard to which the file adheres
    - format: 32-bit signed integer
  - dataLayout
    - example three-dimensional data structure for validating data read orientation
    - size 2 x 3 x 4 (rows x columns x pages)
    - holds increasing integers in column-major order:
      page 1 – [1 3 5 / 2 4 6] page 2 – [7 9 11 / 8 10 12] page 3 – [13 15 17 / 14 16 18] page 4 – [19 21 23 / 20 22 24]
    - format: 32-bit signed integer
  - /ArrayAttributes
    - microphonePositionsM
      - array of microphone coordinates
      - size: microphoneCount x 3
      - units: meters
      - format: floating point
      - note: each row of the array corresponds to a single microphone's coordinate triple as <x,y,z>
    - microphoneCount
      - number of microphones in the microphone array
      - format: 32-bit signed integer

Array Methods HDF5 Time Series File Definitions

- /MetaData
  - /TestAttributes
    - coordinateReference
      - string describing coordinate reference frame and/or origin
      - examples: `array center', `test section entrance'
    - domainBoundsM
      - bounds of the volume containing the source region of interest
      - size: 2 x 3; rows are domain minima and maxima; columns are x-, y- and z-bounds
      - units: meters
      - format: floating point
    - flowType
      - string describing flow field of simulation/measurement
      - examples: `no flow', `uniform flow', `open jet'
    - testDescription
      - string outlining details of simulation/measurement
      - examples: `simulated line source in open jet', `Date; Place; Test point no.; Condition'

Array Methods HDF5 Time Series File Definitions

- /MeasurementData
  - machNumber
    - Mach number of flow field
    - size: 1 x 3, corresponding to x-, y- and z-components
    - format: floating point
  - relativeHumidityPct
    - relative humidity of simulation/measurement
    - units: percent
    - format: floating point
  - speedOfSoundMPerS
    - speed of sound of simulation/measurement
    - units: m/s
    - format: floating point
  - staticPressurePa
    - static pressure of simulation/measurement
    - units: Pa
    - format: floating point
  - staticTemperatureK
    - Static temperature of simulation/measurement
    - units: Kelvin
    - format: floating point

Array Methods HDF5 Time Series File Definitions

- /MicrophoneData
  - microphoneDataPa
    - microphone time series data
    - size: sampleCount x microphoneCount
    - units: Pa
    - format: floating point
    - chunk size: : sampleCount x 1
    - note: each chunk is a single microphone time series
    - note: units are selected to account for microphone sensitivity and any amplifier settings; please remember that some transducers have a negative voltage change for a positive pressure change, and account for this in the conversion of the data
    - sampleCount
      - number of samples in a single microphone time series
      - format: 32-bit signed integer
    - sampleRateHz
      - rate of data acquisition
      - units: samples per second
      - format: floating point

Array Methods HDF5 Time Series File Definitions

- /CsmBuild
  - blockOverlapPts
    - number of samples to overlap blocks when using Welch's modified periodogram method
    - format: 32-bit signed integer
    - note: for no overlap, set to zero; example for 50% overlap would be blockSizePts/2 for even-valued blockSizePts
  - blockSizePts
    - number of samples in a single data block used to construct the CSM
    - format: 32-bit signed integer
    - note: while sampleCount/blockSizePts is not required to be an integer, if it is not this should be stated in the documentation along with how to handle the remainder points
  - fftSign
    - sign of the exponent used to Fourier transform the data ($i\omega t$ vs. $-i\omega t$)
    - format: 32-bit signed integer
    - value: +/- 1
  - frequencyBinCount
    - number of frequency bins in spectral processing
    - format: 32-bit signed integer, maximum value of ceiling(blockSizePts/2)
    - note: sized to span from the dc fft bin to the bin before Nyquist if the maximum value, with the positive-infinity-rounding ceiling operation to allow for odd values of blockSizePts
    - note: this definition means one-sided spectra are used, so all non-dc frequency bins have their power doubled in the final CSM

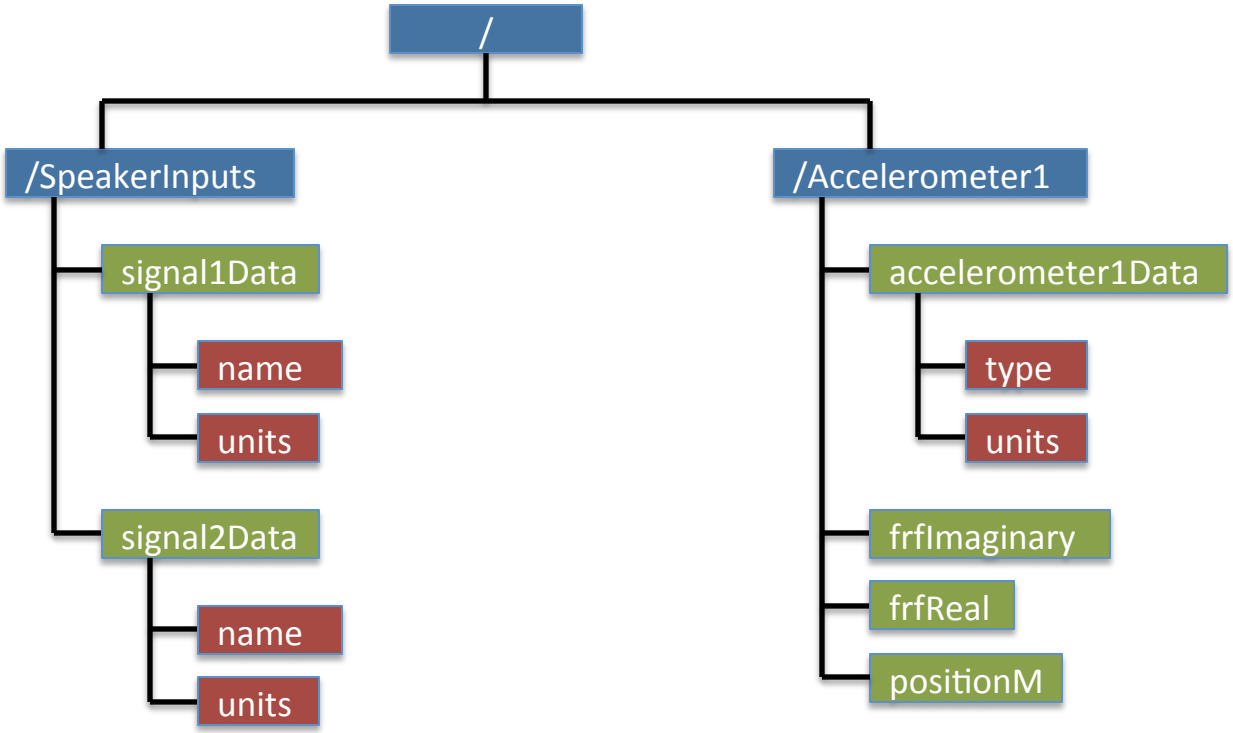
Array Methods HDF5 Time Series File Definitions

- /CsmBuild
  - frfImaginary
    - imaginary component of microphone frequency response functions
    - size: microphoneCount x frequencyBinCount
    - format: floating point
    - chunk size: 1 x frequencyBinCount
    - note: the frequency response function should be normalized, assuming that the transducer sensitivity has already been applied to the data
    - note: defined such that 1/frf corrects the data for the microphone and data acquisition system frequency responses; if no frf is applied, set the imaginary entries to zero
  - frfReal
    - real component of microphone frequency response functions
    - size: microphoneCount x frequencyBinCount
    - format: floating point
    - chunk size: 1 x frequencyBinCount
    - note: the frequency response function should be normalized, assuming that the transducer sensitivity has already been applied to the data
    - note: defined such that 1/frf corrects the data for the microphone and data acquisition system frequency responses; if no frf is applied, set the real entries to unity

Array Methods HDF5 Time Series File Definitions

- /CsmBuild
  - microphoneWeights
    - frequency-independent microphone weighting factors
    - size: microphoneCount x 1
    - format: floating point
    - note: if no weighting factor is used for the data, set all values to unity
  - windowFunction
    - point-by-point window function used in spectral processing
    - size: 1 x blockSizePts
    - format: floating point
    - note: unless specified in the attached documentation, it will be assumed that the broadband window power correction, 1/(sum(windowFunction^2)/blockSizePts), should be applied via multiplication when attempting to reproduce the supplied CSM
    - windowType
      - string describing the type of window function used in spectral processing
      - examples: `boxcar', `hann', `flattop', `custom'

Array Methods HDF5 Example Optional Time Series File



Color Key: /Group attribute dataset

Array Methods HDF5 Example Optional Time Series File

- /SpeakerInputs
  - signal1Data
    - drive signal transmitted to speaker 1
    - size: 1 x sampleCount
    - format: floating point
      - name
        - string giving signal1Data channel/signal name
      - units
        - string giving units for signal1Data
        - example: `volts'
  - signal2Data
    - drive signal transmitted to speaker 2
    - size: 1 x sampleCount
    - format: floating point
      - name
        - string giving signal2Data channel/signal name
      - units
        - string giving units for signal2Data
        - example: `volts'

Array Methods HDF5 Example Optional Time Series File

- /Accelerometer1
  - accelerometer1Data
    - Signal recorded from accelerometer 1
    - size: 3 x sampleCount
    - format: floating point
    - chunk size: 1 x sampleCount
      - type
        - string stating the accelerometer model/type, number of axes
      - units
        - string giving units for accelerometer1Data
        - example: `m/s^2'
  - frfImaginary
    - imaginary component of accelerometer 1 frequency response functions
    - size: 3 x frequencyBinCount
    - format: floating point
    - chunk size: 1 x frequencyBinCount
    - note: the frequency response function should be normalized, assuming that the transducer sensitivity has already been applied to the data
    - note: defined such that 1/frf corrects the data for the accelerometer and data acquisition system frequency responses; if no frf is applied, set the imaginary entries to zero

Array Methods HDF5 Example Optional Time Series File

- /Accelerometer1
  - frfReal
    - real component of accelerometer 1 frequency response functions
    - size: 3 x frequencyBinCount
    - format: floating point
    - chunk size: 1 x frequencyBinCount
    - note: the frequency response function should be normalized, assuming that the transducer sensitivity has already been applied to the data
    - note: defined such that 1/frf corrects the data for the accelerometer and data acquisition system frequency responses; if no frf is applied, set the real entries to unity
  - positionM
    - accelerometer position
    - size: 1 x 3
    - units: meters
    - format: floating point
    - note: coordinate triple of <x,y,z>