MISB MOTION IMAGERY STANDARDS BOARD	**MISB ST 1010.3**
STANDARD	
Generalized Standard Deviation and Correlation Coefficient Metadata	**27 October 2016**

1 Scope

This standard (ST) defines a bit-efficient method for transmitting standard deviation and correlation coefficient data. In support of this method a Standard Deviation and Correlation Coefficient Floating Length Pack (SDCC-FLP) construct is defined. The construct leverages the symmetry of the variance-covariance matrix and the fixed data range of the correlation coefficients to reduce the number of bytes transmitted, in effect compressing the data. This method is, therefore, not extendable to more generic matrix cases. This ST is dependent on context from an invoking standard (or other document), called a *Parent Document*, which provides a list of values (random variables) that can have corresponding standard deviation and correlation coefficients; this list of random variables is called the *Source List*.

2 References

The following references and the references contained therein are normative.

[1] MISB ST 1201.1 Floating Point to Integer Mapping, Feb 2014

[2] ISO/IEC 8825-1:1995 (ITU-T X.690) Information Technology - ASN.1 Encoding Rules - Specification of Basic Encoding Rules (BER), Canonical Encoding Rules (CER), and Distinguished Encoding Rules (DER), 1995

[3] MISB ST 0107.2 Bit and Byte Order for Metadata in Motion Imagery Files and Streams, Feb 2014

[4] SMPTE ST 336:2007 Data Encoding Protocol Using Key-Length-Value,

3 Revision History

Revision	Date	Summary of Changes
ST 1010.3	10/27/2016	• Added support for sets as groups of random variables.

4 Abbreviations and Acronyms

2D	Two Dimensional
BER-OID	Basic Encoding Rules Object identifier

DLP	Defined Length Pack
ECEF	Earth Centered Earth Fixed
FLP	Floating Length Pack
KLV	Key-Length-Value
LS	Local Set
SDCC-FLP	Standard Deviation and Correlation Coefficient - Floating Length Pack
VLP	Variable Length Pack

5 Introduction

A variance-covariance matrix is a NxN square matrix composed of standard deviations and correlation coefficients for a set of N random variables. The techniques described below are a blend of bit efficiency and flexibility for multiple uses of transmitting a variance-covariance matrix to end-user applications.

An example variance-covariance matrix of three random variables is given below in Equation 1.

$$Q = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 \end{bmatrix} \qquad \textbf{Equation 1}$$

This matrix can be divided into two types of parameters: (1) standard deviations; and (2) correlation coefficients. These two parameters are expressed in Equation 2, where σ represents the standard deviations and ρ the correlation coefficients.

$$Q = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{21} & 1 & \rho_{23} \\ \rho_{31} & \rho_{32} & 1 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix} \qquad \textbf{Equation 2}$$

Correlation coefficients are mathematical values describing how two random variables behave relative to one another. They are derived from a variance-covariance matrix produced from a stochastic process, such as least-squares adjustments or Kalman filtering. The correlation coefficient values are bounded between negative-one and positive-one [-1.0, 1.0]. Typical variance-covariance matrices are symmetrical about the main diagonal; for example, the value with subscript (1, 2) is equal to the value with subscript (2, 1). This symmetry reduces the number of values required to reconstruct the full variance-covariance matrix from N^2 to M, an approximate 50% reduction in data as computed in Equation 3.

$$M = \frac{N * (N - 1)}{2} \qquad \textbf{Equation 3}$$

The relationship among correlation coefficients, standard deviation, and covariance (when corresponding values are not zero) is given by Equation 4.

$$\rho_{ij}\sigma_i\sigma_j = \sigma_{ij} \qquad \textbf{Equation 4}$$

This ST defines a bit-efficient method to represent the standard deviation and correlation coefficient values using a Floating Length Pack (FLP), which is a KLV construct used to group KLV items.

6 Standard Deviation and Correlation Coefficient Metadata

This section provides an overview of the encoding method, and formalizes the bit-efficient packaging of standard deviation and correlation coefficient metadata.

6.1 Overview and Definitions

A two-dimensional (2D) standard deviation and correlation coefficient matrix is "encoded" into a linear block of data. A receiver decodes the linear data block re-establishing a 2D standard deviation and correlation coefficient matrix to support further uncertainty propagation. The standard deviation and correlation coefficient matrix can be converted to a variance-covariance matrix to perform subsequent uncertainty propagation.

The standard deviation and correlation coefficient matrix simplifies the transmission process. Figure 1 illustrates the overall flow and aspects of the encoding. The "Source Data" includes the standard deviation and correlation coefficient values, which are encoded and packaged as "Compressed" Data for transmission. Once received, the data is decoded as "Reconstructed Data". The processes of encode/decode are discussed in succeeding sections.

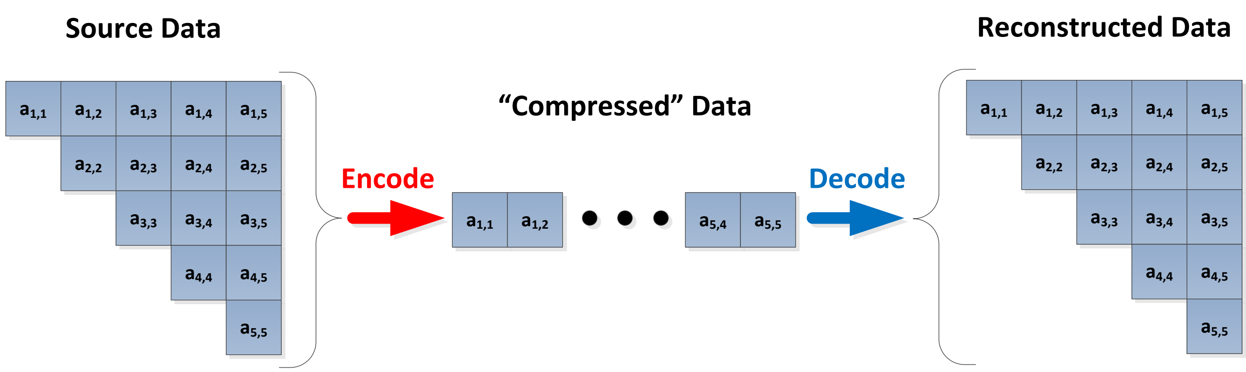

Figure 1: Encode through Decode Process

6.2 Standard Deviation and Correlation Coefficient Matrix Properties

The standard deviation and correlation coefficient matrix is a square symmetric matrix, where the diagonal values of the matrix contain the standard deviation values and the off-diagonal values contain the correlation coefficient values. Because of the symmetry in the correlation coefficients about the diagonal, the lower triangular values are omitted as shown in Figure 2

Each row of the standard deviation and correlation coefficient matrix in Figure 2 represents a random variable compared to a list of other random variables, which are represented by each column. A random variable is created when a measurement or computation is made about a specific phenomenon; for example, a measurement for the ECEF X, Y and Z position of a sensor. The uncertainty propagation parameters for these three items in a 3x3 matrix contain two pieces of information about this measurement: (1) the standard deviation of each individual element (how well the measurement was made – i.e. the error); and (2) the correlation coefficient or relationship of one measurement to another (i.e. the uncertainty in the X measurement is impacted by the uncertainty in the Y measurement).

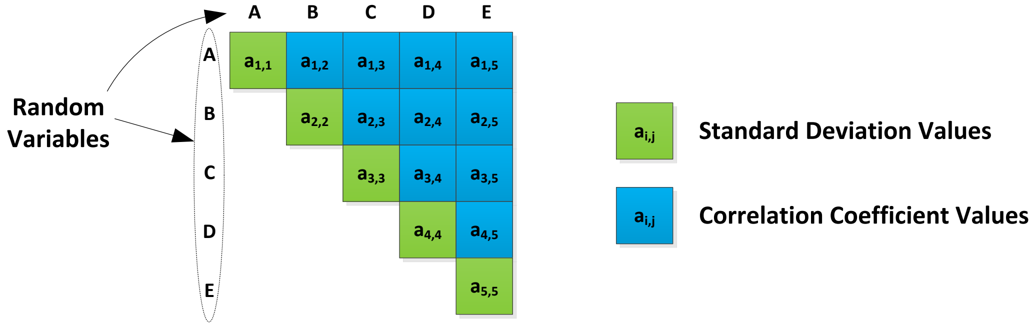

Figure 2: Standard Deviation and Correlation Coefficient Matrix

The matrix in Figure 2 shows an upper triangular portion of the standard deviation and correlation coefficient matrix created through a collection of measurements. There are cases, however, where either individual measurements may not have been made, or the correlation coefficient between two values is unknown (or zero). When one or more measurements are not made, those rows (and corresponding columns) are set to zero, which is the equivalent of removing those rows (and corresponding columns) from the matrix. This in turn reduces the data needed to be sent to a receiver. When a correlation coefficient value between two variables is unknown or zero, that value is eliminated, which causes the matrix to be sparse; this likewise reduces data. Figure 3 illustrates these two cases of the elimination of row/column and a sparse matrix.

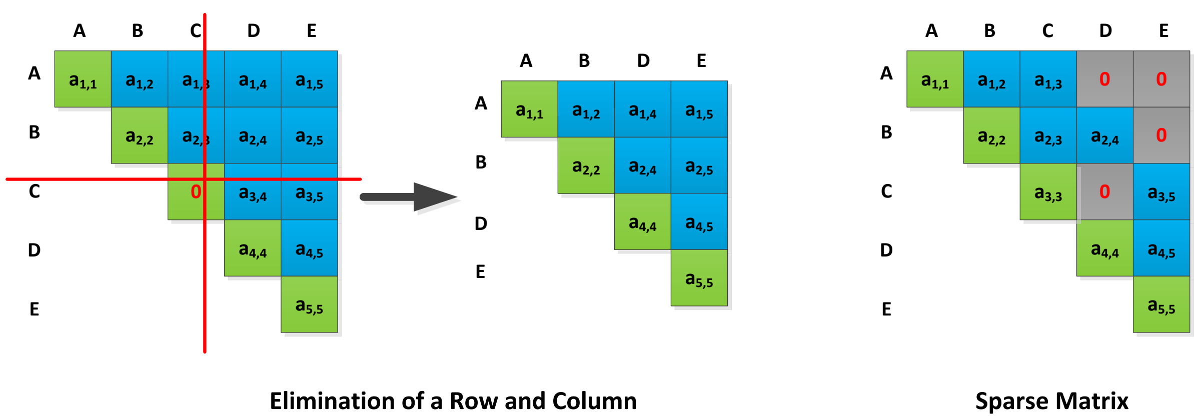

Figure 3: Reduced and Sparse Standard Deviation and Correlation Coefficient Matrix

6.3 Matrix Encoding

6.3.1 Symmetry

One property of a standard deviation and correlation coefficient matrix is symmetry about the diagonal, which means all values in the upper triangular area are equal to those in the lower triangular area. By eliminating the lower triangular portion of the matrix there is a major reduction in the data (bytes) needed for representation, which improves data efficiency. This ST takes advantage of this symmetry by only transmitting the upper triangular area, where the lower triangular values are constructed as needed. Figure 4 shows data in an upper triangular matrix being prepared as a linear block of data for encoding into KLV.

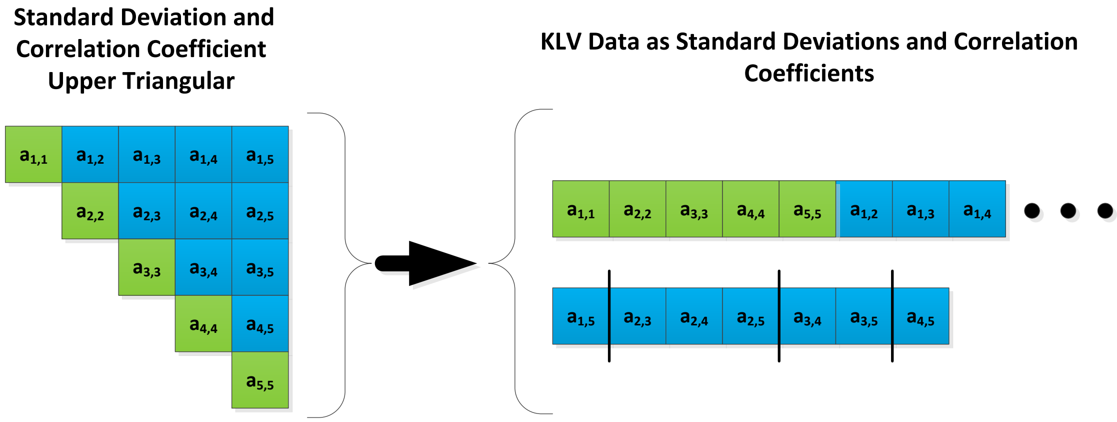

Figure 4: Standard Deviation and Correlation Coefficient Data Prepared for KLV Encoding

6.3.2 Parse Control

The standard deviation and correlation coefficient data are formatted for transmission differently; each will have their own data size and type (float, integer, etc.). Requirements for both are as follows:

Requirement(s)	
ST 1010.2-09	All standard deviation values shall be of a consistent type and size.
ST 1010.1-01	When the Parent Document invoking MISB ST 1010 describes varying byte sizes or data types for the standard deviation parameters, the greatest or most conservative value shall be used for encoding to preserve precision.
ST 1010.2-10	All correlation coefficient values shall be of a consistent type and size.
ST 1010.1-02	When the Parent Document invoking MISB ST 1010 describes varying byte sizes for the correlation parameters, the greatest or most conservative value shall be used for encoding to preserve precision.

As an example, consider a Parent Document which specifies ECEF X, Y, and Z positions, along with standard deviations to describe the uncertainty of these values. The Parent Document will define the variable size and type of the standard deviation values and correlations coefficient values. Assume some standard deviation values are represented as four-byte floating point values, and others as two-byte MISB ST 1201 [1] mapped integer values. As all standard deviation values in a SDCC-FLP are required to have the same length, the two-byte integer values are coded as four-byte floating point values; this preserves the precision of all the values (see Figure 5).

Likewise, the correlation coefficient values are required to have the same number of bytes for each value within a FLP. Since the range for a correlation coefficient value is [-1.0, 1.0], the correlation coefficient values should be mapped into integer values according to ST 1201 to potentially reduce the number of bytes for each correlation coefficient value. Continuing with the example, assume all the correlation coefficient values for the ECEF X, Y, and Z positions are represented using two-byte ST 1201 values. The complete stream of data values is as shown in Figure 5.

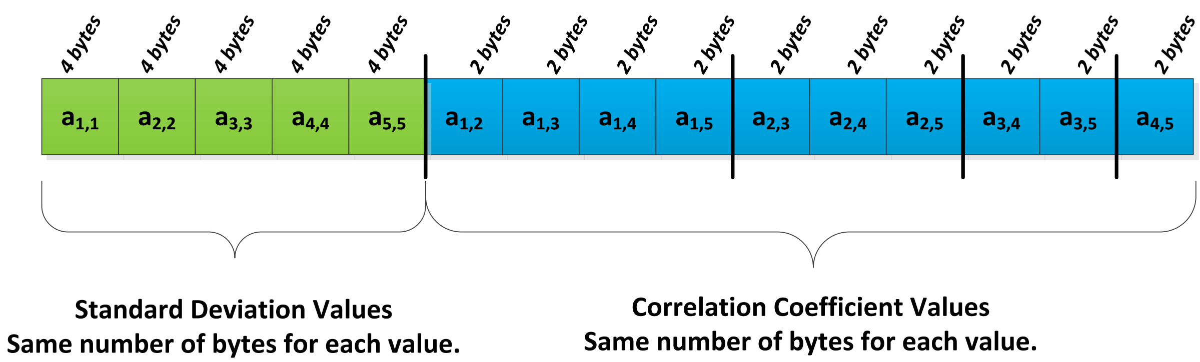

Figure 5: Example: Variance/Covariance Values Data Type Consistency

To interpret data such as shown in Figure 5 and parse the correct number of bytes, the data size and type for both the standard deviation (S_{len}) and correlation coefficient (C_{len}) values are included in the data stream. The data size and type information are encoded into one or more bytes called *Parse Control*. The Parse Control information precedes the standard deviation and correlation coefficient data in the FLP. Parse Control can have a different length depending on the application, so the encoding method of Basic Encoding Rules Object identifier (BER-OID) [2] byte extension is used. The eighth bit in each byte signals if there are more bytes to follow; a value of 1 indicates there is another following byte, while a value of zero indicates this is the last byte.

Parse Control offers two modes: Mode 1 and Mode 2. Mode 1 maintains backward compatibility when parsing data based on version 1 of this standard; this mode has limited lengths (S_{len} and C_{len}) and the value types are defined by the Parent Document. Mode 2 adds additional length capabilities and provides for runtime determination of data types, either ST 1204 or IEEE floating point values.

6.3.2.1 Mode 1

With Mode 1 Parse Control, a standard deviation value's length can range from 0 to 7 bytes, and the type for all values are either ST 1201 mapped floating point values or IEEE floating point values as specified by the Parent Document. The length of the correlation coefficient values can also range from 0 to 7 bytes, but the type is restricted to be a ST 1201 mapped floating point value. Mode 1 is encoded as one byte as shown in Figure 6 and described in Table 1.

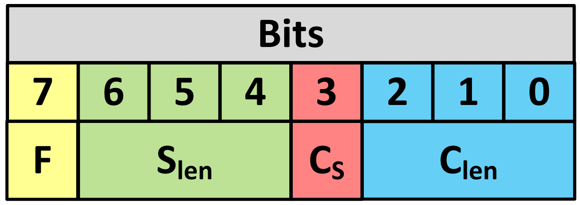

Figure 6: Mode 1 – Parse Control Bit Assignments

Table 1: Mode 1 Parse Control Bits

Bit(s)	Color	Value or Variable	Description
7	Yellow	F	"Final" byte indicator: 0 indicates last byte; 1 indicates more Parse Control bytes to follow. For Mode 1 this bit is always set to zero.
4-6	Green	S_{len}	Three-bit integer defining the number of bytes used for the standard deviation values. Bit six is the most significant bit. Provides for a data size of up to seven bytes maximum. A value of zero indicates there are no standard deviation values.
3	Red	C_S	A flag for spare matrix selection, which is discussed further in Section 6.3.4. A value of one indicates a sparse matrix is used, and a value of zero indicates a sparse matrix is not used.
0-2	Blue	C_{len}	Three-bit integer defining the number of bytes used for the correlation coefficients. Bit two is the most significant bit. Provides for a data size of up to seven bytes maximum. A value of zero indicates there are no correlation coefficient values.

Figure 7 illustrates an example of a Mode 1 Parse Control, which is specified using one byte. The value, 0x4B in hex, signifies that standard deviation values are four bytes, the correlation coefficients are stored using a sparse method, and the correlation coefficients are three bytes in length.

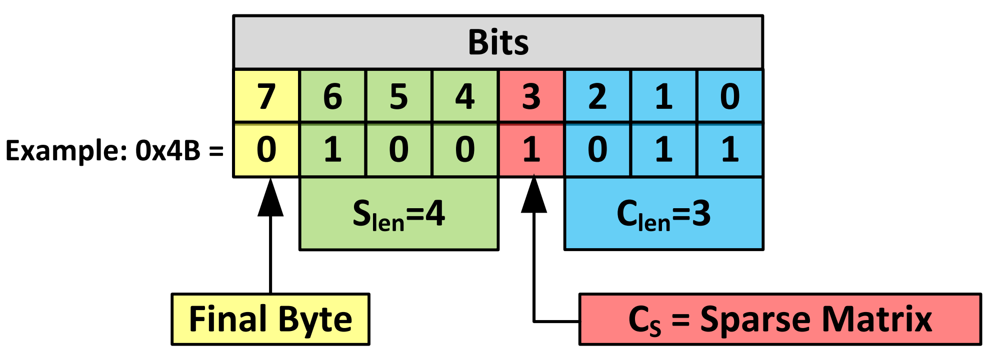

Figure 7: Mode 1 Parse Control Example

6.3.2.2 Mode 2

With Mode 2 Parse Control, a standard deviation value's length can range from 0 to 15 bytes, and the type for all values are either ST 1201 mapped floating point values or IEEE floating point values. The type is specified within the Mode 2 Parse Control. The length of the correlation coefficients can also range from 0 to 15 bytes, and are either ST 1201 mapped floating point values or IEEE floating point values, which is specified within the Mode 2 Parse Control.

Mode 2 is encoded as two bytes as shown in Figure 8 and described in Table 2. Mode 2 is defined by the following bits: the mode (bit 7 yellow) in Byte 1 is set to one to indicate more bytes follow; in this case Byte 2. The mode selection (bit 7 yellow) in Byte 2 is set to zero to indicate this is the last byte.

Byte 1								**Byte 2**							
7	**6**	**5**	**4**	**3**	**2**	**1**	**0**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**
F_1	R	C_S	C_f	C_{len}				F_2	R		S_f	S_{len}			

Figure 8: Mode 2 – Parse Control Bit Assignments

Table 2: Mode 2 Parse Control Bits

	Bit(s)	Color	Value or Variable	Description
Byte 1	7	Yellow	F_1	"Final" byte indicator: 0 indicates last byte; 1 indicates more bytes to follow. For Mode 2 this bit is set to one in Byte 1.
	6	Grey	R	This bit is reserved and always set to zero for Mode 2.
	5	Red	C_S	A flag for spare matrix selection, which is discussed further in Section 6.3.4. A value of one indicates that a sparse matrix is used, and a value of zero indicates that a sparse matrix is not used.
	4	Purple	C_f	Correlation coefficients format. A value of 0 indicates IEEE floating point format is used. A value of 1 indicates ST 1201 format is used with a min value of -1 and max value of +1.
	0-3	Blue	C_{len}	Four-bit integer defining the number of bytes used for the correlation coefficients. Bit three is the most significant bit. Provides for a data size of up to 15 bytes maximum. A value of zero indicates there are no correlation coefficient values.
Byte 2	7	Yellow	F_2	"Final" byte indicator: 0 indicates last byte; 1 indicates more bytes to follow. For Mode 2 this bit is set to zero in Byte 2.
	5,6	Grey	R	These bits are reserved and always set to zero for Mode 2.

	4	Purple	S_f	Standard deviation format. A value of zero indicates IEEE floating point format is used. A value of 1 indicates ST 1201 format is used with the min and max values defined by the Parent Document.
	0-3	Green	S_{len}	Four-bit integer defining the number of bytes used for the Standard Deviation values. Bit three is the most significant bit. Provides for a data size of up to 15 bytes maximum. A value of zero indicates there are no standard deviation values.

Figure 9 illustrates an example of Mode 2 Parse Control, which is specified using two bytes. The value in hex is 0xB308, signifying that the correlation coefficient values are stored using a sparse method each represented as a mapped ST 1201 value using three bytes, and the standard deviation values are represented as eight-byte IEEE floating point values (double precision).

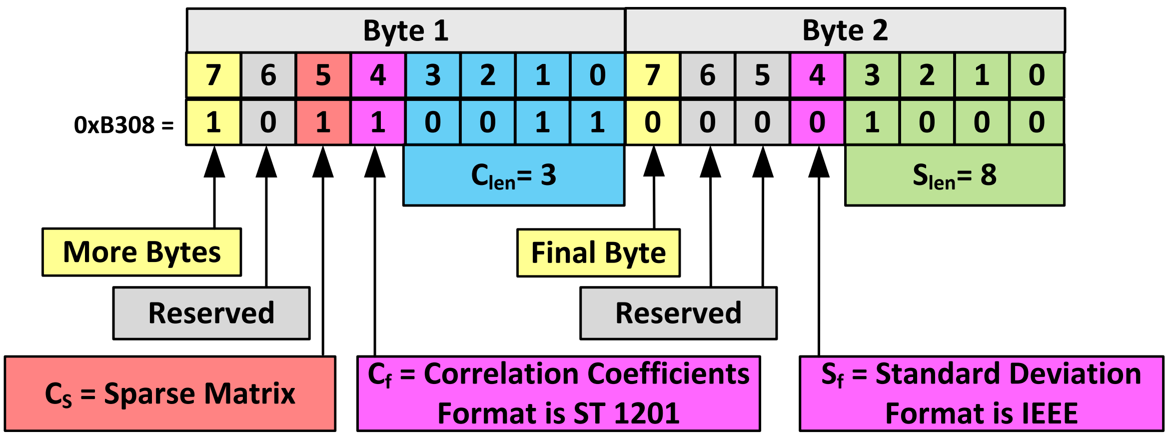

Figure 9: Mode 2 Parse Control Example

The values for C_f and S_f indicate whether the data values are formatted in IEEE floating point or MISB ST 1201. When correlation values use ST 1201, the minimum value is -1 and the maximum value is +1. For standard deviation values, the Parent Document needs to supply the minimum and maximum values used.

6.3.2.3 Data Sizes

A data size (S_{len} or C_{len}) with a value of zero means there are no data of that type (standard deviation or correlation coefficient values) in the data set. For example, if the correlation coefficient data length C_{len} is zero, then only standard deviation data is present. Figure 10 illustrates the addition of the Parse Control; the top graphic shows length information (S_{len} = 4) for standard deviation data (4 bytes per value) and length information (C_{len} = 2) for correlation coefficient data (2 bytes per value). All corresponding data for standard deviation and correlation coefficient values follows. The bottom graphic shows length information (S_{len} = 4) for standard deviation data (4 bytes per value) and length information (C_{len} = 0) for correlation coefficient data (0 bytes per value). In this bottom case, since C_{len} = 0, no correlation coefficient data follows.

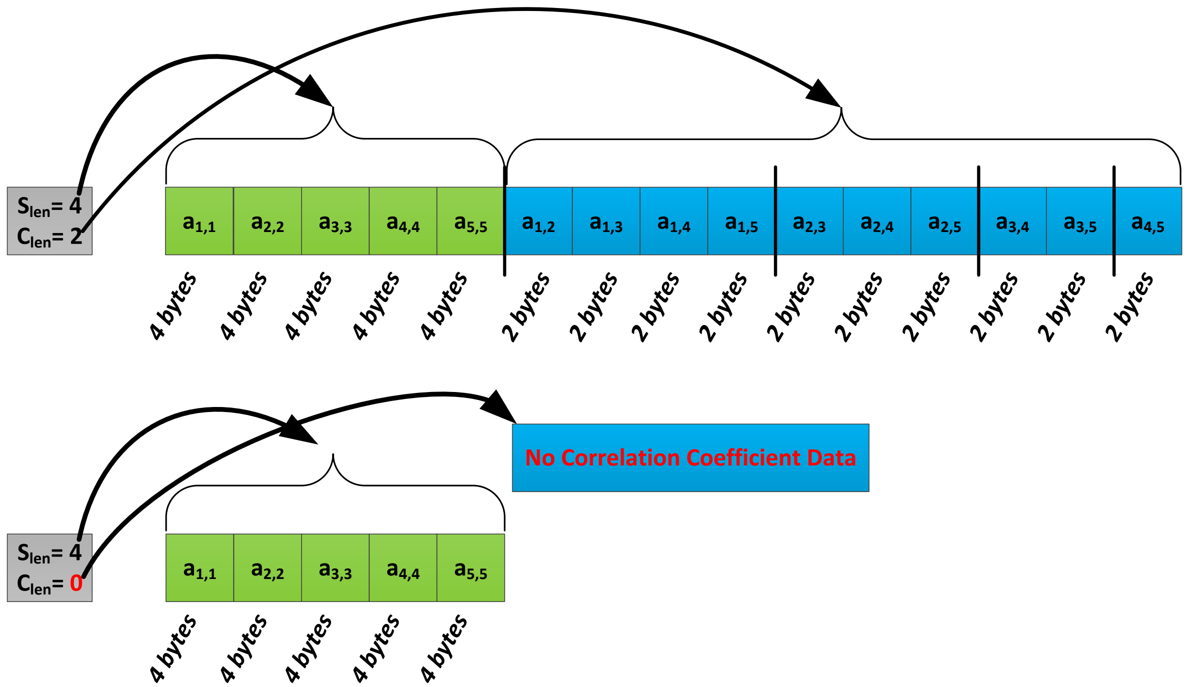

Figure 10: Standard Deviation and Correlation Coefficient Data with Parse Control Information

6.3.3 Matrix Size

In order to parse and reconstruct a standard deviation and correlation coefficient matrix, the size of the matrix is needed. The Matrix Size is included in the SDCC-FLP as a BER-OID encoded value. This affords defining matrix sizes of unlimited dimension. Matrix dimensions up to 127 are described using one byte for Matrix Size (bit 7 = 0). Greater sizes are described using additional bytes as controlled by bit 7 = 1 with bit 7 = 0 in the terminating byte. Figure 11 shows the relation of Matrix Size N to the lengths of each row.

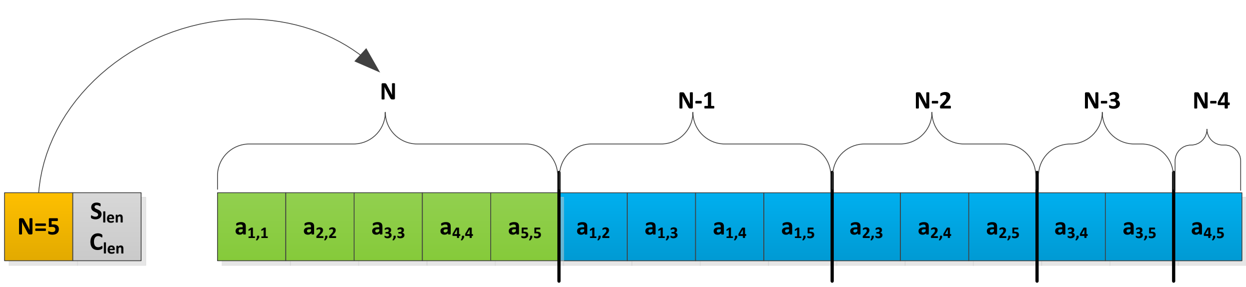

Figure 11: Standard Deviation and Correlation Coefficient Data with Parse Control and Matrix Size Information

6.3.4 Bit Vector (Sparse Representation)

There are cases where correlation coefficient values are unknown or zero (i.e. sparse) – see Section 8 for more information. To provide for sparse matrices, a *Bit Vector,* which is optional, signals whether a value is present or not. The inclusion of the Bit Vector is controlled by the sparse matrix **(**C_S) bit within Parse Control (see Figure 6 and Figure 8).

Requirement	
ST 1010.2-11	The Bit Vector shall only be used to convey a sparse matrix (i.e. a Bit Vector cannot convey a full matrix by populating with all ones.)

Table 3: Bit Vector Usage as a Function of C_{len} and C_S

C_{len}	C_S	Correlation Coefficient Values	Bit Vector
0	X	None present in the data	Not Needed
> 0	0	Full set present in the data	Not Needed
> 0	1	Sparse set present in the data	Denotes where values present

The Bit Vector can only be used if the correlation coefficient values form a sparse matrix. When present, it is the same order and length as the correlation coefficient values. The number of bytes constituting the Bit Vector is computed by associating one bit to each represented correlation coefficient value, and then rounding the number up to the nearest byte. The Bit Vector starts with the most significant bit of the most significant byte, and ends with zero padding in the least significant bits of the last byte (see top graphic in Figure 12). The Bit Vector only applies to correlation coefficient data, and thus, the length of the Bit Vector depends on the maximum number of correlation coefficient values i.e. **M** (see Equation 3). Equation 5 is used to compute the length, in bytes, of the Bit Vector.

$$V_{bytes} = \left\lceil \frac{C_s M}{8} \right\rceil = \left\lceil \frac{C_s \left(\frac{N(N-1)}{2} \right)}{8} \right\rceil = \left\lceil \frac{C_s N(N-1)}{16} \right\rceil \qquad \textbf{Equation 5}$$

Where: $\lceil y \rceil$=ceiling(y) (i.e. round up); N = Matrix Size

As an example, Figure 12 shows a full matrix (N= 5) with several correlation coefficient values not populated. Values to be included are indicated with a “1” and values not included with a “0”. This collection of ones and zeros comprise the Bit Vector. The calculation shows two bytes are needed to convey the Bit Vector. The bottom graphic shows the resulting block of data transmitted.

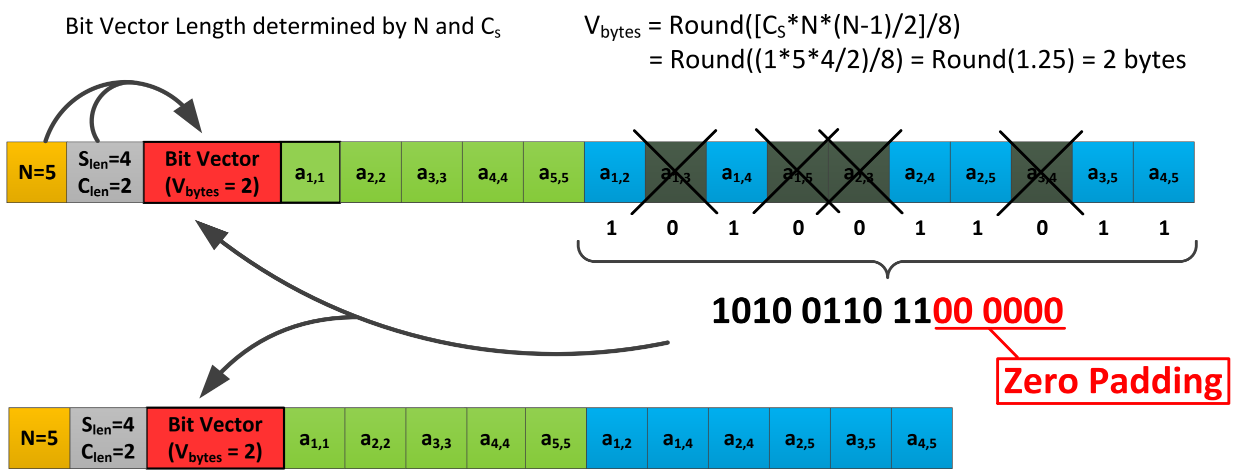

Figure 12: Standard Deviation and Correlation Coefficient Data with Parse Control, Matrix Size, and Bit Vector Information

In the case where standard deviation data is not accompanied with correlation coefficient data, C_{len} is set to zero, C_s is set to zero, and the Bit Vector is not necessary. An example of this is shown in Figure 13.

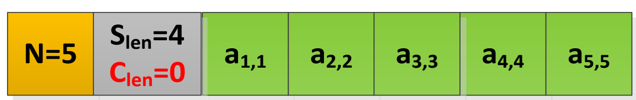

Figure 13: Standard Deviation Data Without Bit Vector

These examples demonstrate the flexibility and the efficiency of this method. See Appendix A for a cost analysis of when it is beneficial to use a Bit Vector.

7 Standard Deviation and Correlation Coefficient FLP (SDCC-FLP)

The SDCC-FLP is composed of a multitude of items combined into one block of KLV binary data. Five elements, in the order specified in Table 4, comprise the FLP: Matrix Size, Parse Control, Bit Vector, Standard Deviation values, and the Correlation Coefficient values.

When processing is performed in support of the Parent Document's metadata needs, it is called *Parent Processing*. All of the elements in the SDCC-FLP are "computed" at run-time by Parent Processing. The Parent Processing specifies the Matrix Size N, which is less than or equal to the number of random variables in the Parent Documents Source List.

The SDCC-FLP affords efficiencies when a matrix is sparse, or lacks correlation values; attention to how the five elements in the SDCC-FLP are populated can reduce the data required to transmit standard deviation and correlation coefficient information. For instance, random variables may be removed by reordering the data set parameters, which effectively equates a row and column for a value not containing a standard deviation to zero, thereby reducing bytes (see example in Section 8.3. Parse Control and Bit Vector are set by the Parent Processing to characterize and define the contents of data present in the SDCC-FLP.

Multiple SDCC-FLPs are allowed within a metadata stream. Each SDCC-FLP contains a subset of the Source List specified in the Parent Document. Option 3 in the example Section 8 illustrates this use case.

Table 4: SDCC-FLP

FLP Key		Name	

`06 0E 2B 34 02 05 01 01`
`0E 01 03 03 21 00 00 00`
(CRC 64882)
		Standard Deviation and Correlation Coefficient FLP	

Element Name	**Type**	**Size**	**Description**
Matrix Size (Required)	BER OID Integer	Variable	Describes the dimension, N, of the matrix of standard deviation and correlation coefficient values to be transmitted. This value is a BER-OID encoded count of the number of rows or columns in a square matrix.
Parse Control (Required)	Binary	1 Byte (Mode 1)	Contains three values (C_S, S_{len}, C_{len}) encoded as a single byte. C_S is a single bit indicating if the correlation coefficient values are sparsely represented or not; if the C_S bit is set to true (1) then the Bit Vector (next element) is included in this data block. The S_{len} and C_{len} values indicate the number of bytes used for the standard deviation values and the correlation coefficient values, respectively. If S_{len} is zero, then there are no standard deviation values in the data block. If C_{len} is zero, then there are no correlation coefficient values in the data block. See Figure 6 for the bit assignments.
		2 Bytes (Mode 2)	Contains five values (C_s, S_{len}, C_{len}, S_f, C_f) encoded into two bytes. C_S is a single bit indicating if the correlation coefficient values are sparsely represented or not; if the C_S bit is set to true (1) then the Bit Vector (next element) is included in this data block. The S_{len} and C_{len} values indicate the number of bytes used for the standard deviation values and the correlation coefficient values, respectively. If S_{len} is zero, then there are no standard deviation values in the data block. If C_{len} is zero, then there are no correlation coefficient values in the data block. The S_f and C_f values indicate if the standard deviation and correlation coefficients values are encoded using the method from MISB ST 1201 [1] or IEEE floating point values. See Figure 8 for the bit assignments.
Bit Vector (Optional)	Array of Bits	Variable	An optional group of bits denoting which values of the correlation coefficients are being transmitted in the FLP. The size of this group is computed using Equation 5. The Bit Vector along with the Matrix Size is used to determine where values populate the two-dimensional standard deviation and correlation coefficient array. The existence of this data is controlled by the values contained in C_S and C_{len}. (See Table 3).
Standard Deviation Values (Conditional)	Array of Values	Variable (Mode 1)	A list of values where each value has a byte size specified by S_{len}. The number of values is dependent on the Matrix Size, specified by the Parent Processing. Each value is a MISB ST 1201 [1] mapped floating point value, or IEEE floating point value as specified by the Parent Document.

		Variable (Mode 2)	A list of values where each value has a byte size and type specified by S_{len} and S_f respectively. The number of values is depended on the Matrix Size, specified by the Parent Processing. Each value is a MISB ST 1201 [1] mapped floating point value, or IEEE floating point value as specified by the Parent Document.
Correlation Coefficient Values (Conditional)	Array of Integer Mapped Values	Variable (Mode 1)	A list of values where each value has a byte type and size specified by C_{len}. The number of values is depended on the Matrix Size and the bits in the Bit Vector. Each value is a floating point-to-integer mapped value using MISB ST 1201 [1] for mapping the values. The range of each value is [-1.0, 1.0], so the mapping function, as specified by ST 1201, is IMAPB(-1.0, 1.0, C_{len}).
		Variable (Mode 2)	A list of values where each value has a byte size and type specified by C_{len} and C_f respectively. The number of values is dependent on the Matrix Size and the bits in the Bit Vector. When each value is represented using MISB ST 1201 [1] the range of each value is [-1.0, 1.0], so the mapping function, as specified by ST 1201, is IMAPB(-1.0, 1.0, C_{len}).

The following are requirements when implementing the SDCC-FLP.

Requirement(s)	
ST 1010.1-04	The SDCC-FLP shall contain all the required parameters in MISB ST 1010 Table 4 SDCC-FLP.
ST 1010.2-12	The SDCC-FLP shall contain one or both conditional parameters in MISB ST 1010 Table 4 SDCC-FLP.
ST 1010.1-05	When a Parent Document requires standard deviation and/or correlation coefficient information, the Parent Processing shall populate the five parameters described in the SDCC-FLP in accordance with MISB ST 1010 Table 4.
ST 1010.1-08	All metadata shall be expressed in accordance with MISB ST 0107 [3].

8 Invoking ST 1010

The SDCC-FLP is but one item within a Parent Document's metadata items. There are three types of SDCC-FLP KLV constructs a Parent Document can use: Defined Length Pack (DLP), Variable Length Pack (VLP) and Local Set (LS). In all three constructs the Parent Document defines a Source List (the random variables) and one or more SDCC-FLPs. In defining its data structure, a Parent Document will have a list of KLV items. From this list the Parent Document will indicate those items which are random variables (or sets of random variables); these items become the Source List (See Figure 14).

Figure 14 illustrates an example of items (Item 10 - Item 17) in a KLV construct. A Source List includes five random variables (in blue). The succeeding SDCC-FLP states the number of preceding random variables in the Source List to be used for the SDCC-FLP. Items (in grey) above the Source List (in blue) and below the SDCC-FLP (in green) are independent of the SDCC-FLP.

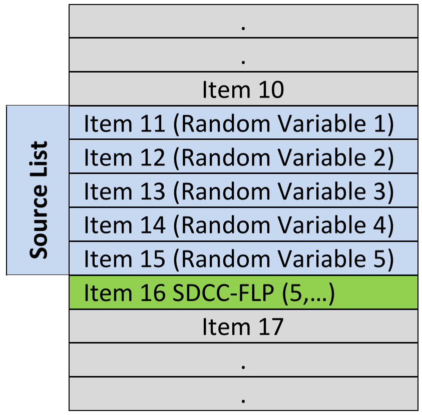

Figure 14: Source List and SDCC-FLP Example

In some cases, not all items in the Source List are available. For example, while there may be 30 possible items in the Source List which could be represented, perhaps only 10 are. In this case not all items on the Source List need be included. To account for this possibility, a *Refined Source List*, which is a subset of the Source List, identifies only those items that are represented. At run-time the Parent Processing creates this Refined Source List. The random variables in the Refined Source List need to match the order of the values in the SDCC-FLP.

Requirement	
ST 1010.1-06	At runtime the ordering of the values in a SDCC-FLP shall correspond to the same ordering of the random variables in the Refined Source List.

8.1 KLV Constructs and SDCC-FLP

When using a SDCC-FLP the different types of KLV constructs (DLP, VLP and LS) each afford a different degree of flexibility, as discussed below. Within each of the KLV constructs, KLV items are used to define the Source List; items can be standalone or a group as discussed in Section 8.1.4.

8.1.1 Defined Length Pack (DLP)

In using a DLP, SMPTE ST 336 [4] rules state that each item in the DLP need to be in a pre-defined order with a pre-defined size. With the DLP the Parent Document will:

1) Define the Source List's random variables.
2) Define the order of the Source List's random variables.
3) Include the SDCC-FLP after the Source List's random variables.
4) Define the values for all SDCC-FLP parameters (N, C_s, S_{len}, C_{len}, S_f, C_f).

With DLP the Refined Source List is <u>always the same</u> as the Parent Document's Source List. Figure 15 illustrates a DLP with the Parent Document defined Source List on the left and the resulting DLP at runtime.

Parent Document		DLP Representation at Run time Value	
	.	.	
	.	.	
	Item 10	Item 10	
Source List	Item 11 (Random Variable 1)	Item 11 (Random Variable 1)	Refined Source List
	Item 12 (Random Variable 2)	Item 12 (Random Variable 2)	
	Item 13 (Random Variable 3)	Item 13 (Random Variable 3)	
	Item 14 (Random Variable 4)	Item 14 (Random Variable 4)	
	Item 15 (Random Variable 5)	Item 15 (Random Variable 5)	
	Item 16 SDCC-FLP (5,…)	SDCC-FLP (5,…)	
	Item 17	Item 17	
	.	.	
	.	.	

Figure 15: Example DLP: note that Refined Source List = Source List

8.1.2 Variable Length Pack (VLP)

In using a VLP, SMPTE ST 336 [4] rules state that all items in the VLP must be included and the length of each item can be changed; however, the order must remain as defined in the Parent Document. With the VLP the Parent Document will:

1) Define the Source List's random variables.
2) Define the order of the Source List's random variables.
3) Include the SDCC-FLP after the Source List's random variables.

The SDCC-FLP parameters (N, C_s, S_{len}, C_{len}, S_f, C_f) can be adjusted at run-time, so the Refined Source List can be a subset of the Source List random variables and any value representations (i.e. IMAP or IEEE) can be used. The order of random variables in the SDCC-FLP must match the order in the Parent Document, and only the random variables immediately preceding the SDCC-FLP are used for the Refined Source List.

Figure 16 illustrates a VLP with the Parent Document defined Source List on the left and the resulting VLP on the right. At run-time, when building the SDCC-FLP data, a decision to only include three of the five random variables (highlighted in yellow) can be made by setting the "N" value of the SDCC-VLP to three. This limits the Refined Source List to be only the three random variables above the SDCC-VLP so, the SDCC-FLP will only contain standard deviation and correlation coefficient values for the preceding three items (Item 13, Item 14 and Item 15 - in that order).

Parent Document		VLP Representation at Run time		
		Len	Value	
	.	.	.	
	.	.	.	
	Item 10	6	Item 10	
Source List	Item 11 (Random Variable 1)	4	Item 11 (Random Variable 1)	
Source List	Item 12 (Random Variable 2)	4	Item 12 (Random Variable 2)	
Source List	Item 13 (Random Variable 3)	4	Item 13 (Random Variable 3)	Refined Source List
Source List	Item 14 (Random Variable 4)	4	Item 14 (Random Variable 4)	Refined Source List
Source List	Item 15 (Random Variable 5)	4	Item 15 (Random Variable 5)	Refined Source List
	Item 16 SDCC-FLP (5,…)	25	SDCC-FLP (3 ,…)	
	Item 17	1	Item 17	
	.	.	.	
	.	.	.	

Figure 16: Example VLP: note Refined Source List ≤ Source List

8.1.3 Local Set (LS)

When the Parent Document uses the LS construct, the items in the LS can be of different sizes, in any order, and do not have to be included at run-time; this provides the most flexibility when using the SDCC-FLP. The Parent Document defines the random variables in the Source List. At runtime, the SDCC-FLP parameters (N, C_s, S_{len}, C_{len}, S_f, C_f) can be adjusted along with reordering the random variables in the LS, thereby reducing the SDCC-FLP dataset size.

The Refined Source List is constructed by removing and reordering the random variables, along with their value representations (i.e. IMAP vs IEEE). Unlike the DLP and VLP constructs, SMPTE ST 336 [4] rules do not enforce an ordering of elements in Local Sets at run time; however, when using ST 1010 the order of the Refined Source List's random variables in the Local Set must match the order of the values in the SDCC-FLP, and the SDCC-FLP must immediately follow the Refined Source List.

Requirement	
ST 1010.2-13	When using a Local Set SDCC-FLP, at runtime the SDCC-FLP shall be preceded by the items from the Refined Source List.

Figure 17 illustrates an example LS with the Parent Document defined Source List on the left and the resulting LS on the right. At runtime the Refined Source List is created by reordering and/or removing unneeded random variables, which reduces the SDCC-FLP dataset size (highlighted in yellow). Item 15 is not included and the SDCC-FLP is reduced in size by setting the N value of the SDCC-FLP to three. The SDCC-FLP will then only contain standard deviation and correlation coefficients for the preceding three items (Item 12, Item 13 and Item 11 - in that order).

Parent Document		LS Representation at Run time			
		Tag	Len	Value	
	.	.	.	.	
	.	.	.	.	
	Item 10	.	.	.	
Source List	Item 11 (Random Variable 1)	.	.	.	
Source List	Item 12 (Random Variable 2)	14	4	Item 14 (Random Variable 4)	
Source List	Item 13 (Random Variable 3)	12	4	Item 12 (Random Variable 2)	Refined Source List
Source List	Item 14 (Random Variable 4)	13	4	Item 13 (Random Variable 3)	Refined Source List
Source List	Item 15 (Random Variable 5)	11	4	Item 11 (Random Variable 1)	Refined Source List
	Item 16 SDCC-FLP (5,...)	16	25	SDCC-FLP (3,...)	
	Item 17	.	.	.	
	.	.	.	.	
	.	.	.	.	

Figure 17: Example LS: note removal/renumbering of Random Variables in Source List into Refined Source List

An additional capability afforded by the Local Set construct is the ability to repeat the SDCC-FLP item multiple times within one Local Set. A SDCC-FLP item can be repeated for a different list of random variables taken from the Source List. In this case, each SDCC-FLP is preceded by its own Refined Source List. When two or more Refined Source Lists contain the same random variable care needs to be taken to prevent conflicting standard deviation or correlation coefficient values within the SDCC-FLPs. Figure 18 illustrates a SDCC-FLP used multiple times within a single LS. Refer to Section 8.3 for a discussion on when to use multiple SDCC-FLPs.

Parent Document		LS Representation at Run time			
		Tag	Len	Value	
	.	.	.	.	
Source List	Item 9 (Random Variable 1)	15	4	Item 15 (Random Variable 7)	Refined Source List 1
Source List	Item 10 (Random Variable 2)	9	4	Item 9 (Random Variable 1)	Refined Source List 1
Source List	Item 11 (Random Variable 3)	10.	4	Item 10 (Random Variable 2)	Refined Source List 1
Source List	Item 12 (Random Variable 4)	16	25	SDCC-FLP (3,...)	
Source List	Item 13 (Random Variable 5)	12	4	Item 12 (Random Variable 4)	Refined Source List 2
Source List	Item 14 (Random Variable 6)	13	4	Item 13 (Random Variable 5)	Refined Source List 2
Source List	Item 15 (Random Variable 7)	11	4	Item 11 (Random Variable 3)	Refined Source List 2
	Item 16 - SDCC-FLP (7,...)	16	25	SDCC-FLP (3,...)	
	.	.	.	.	
	.	.	.	.	
	.	.	.	.	

Figure 18: Example LS with multiple SDCC-FLP

8.1.4 KLV Items

There are two types of KLV items used within KLV constructs: individual elements and element groups. Individual elements are directly identified as random values in the Source List of the parent document, as shown in Figure 14.

When a KLV item is a group of elements (i.e. and embedded DLV, VLP or LS), the KLV construct and the element groups form a hierarchy of KLV data, as illustrated in Figure 19. Here the Top Level KLV Construct contains a SDCC-FLP, and a Sub-Level KLV Construct, which itself has a Sub-Sub-Level KLV Construct. Care must be taken to map the hierarchy of elements into the Refined Source List. The element groups have defining documentation (i.e. their own standard or section within a standard), and must include its own Source List if the element group is to be used for the SDCC-FLP. The items in the element groups may or may not be included in Source List, as illustrated with Item 1 in the Sub-Level KLV Construct of the figure. For example, Item 1 could be a pack's version number or state information, which would not be included as a random variable in the Source List. When the elements are not a part of the Source List, they are ignored when creating the Refined Source List.

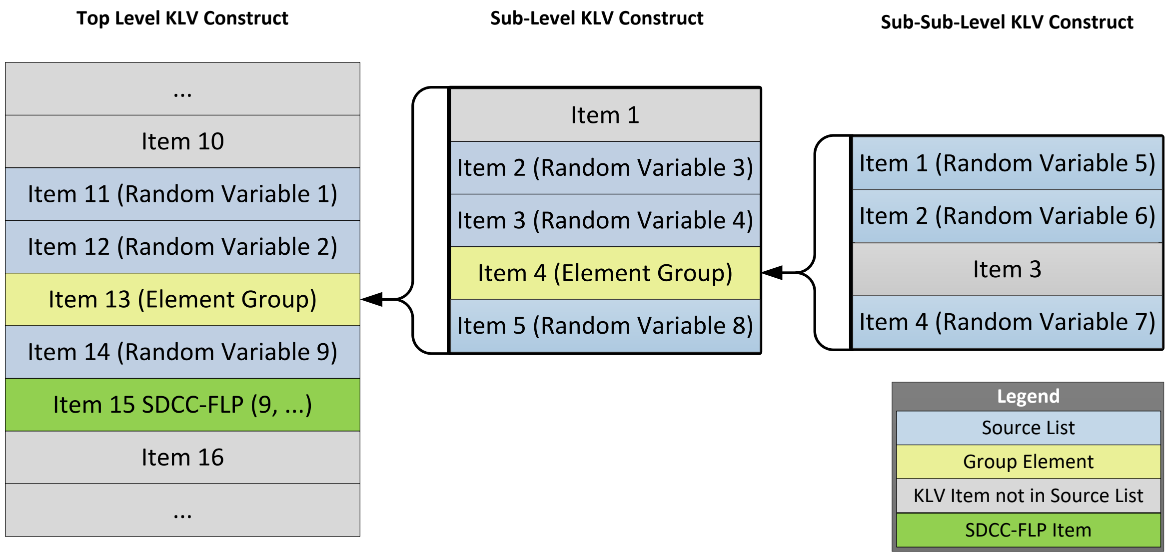

Figure 19: Hierarchical KLV Example

The order of the elements in the hierarchical sub-level KLV constructs, in combination with the elements from the top level KLV construct, must match the order of the elements in the SDCC-FLP.

8.2 SDCC-FLP Preparation

When preparing to build a SDCC-FLP dataset the following run-time process is recommended when the Parent Processing is using either the VLP or LS:

1. Determine the list of random variables in the Source List which require uncertainty information. Define the Refined Source List and the value N.
    a. For a LS determine the order of the random variables.
2. Select the most appropriate standard deviation and correlation coefficient data type, either IMAP or IEEE float, defining S_f and S_c.
3. Select the standard deviation and correlation coefficient size to support the implementer's precision requirements, defining S_{len} and C_{len}.

4. Determine if correlation coefficients are sparse, defining C_s and the Bit Vector.
5. Encode MISB ST 1010 SDCC-FLP, per specification of this document.

8.3 Grouping Source Lists and SDCC-FLP

ST 1010 can be invoked multiple times from a Parent Document, which enables different methods of grouping Source Lists, potentially providing some bandwidth savings. This grouping occurs on a case-by-case basis, and the following example discusses different options.

Assume a Parent Document uses a Local Set construct with twelve random variables and only six of the twelve random variables require standard deviations and correlation coefficients. Furthermore, of the six random variables one set of three (Sensor Position X, Y, Z) do not have correlations with the other set of three (Sensor Heading, Pitch and Roll). There are three options for invoking ST 1010: (1) placing the SDCC-FLP at the end of the twelve random variables; (2) ordering the random variables into one group; and (3) further ordering into multiple groups (i.e. multiple instances of the SDCC-FLP).

8.3.1.1 Option 1

Option 1 requires no rearranging of the tag IDs present in the data set's KLV stream. At the end of the twelfth tag number, the SDCC-FLP is inserted into the KLV stream (see Table 5).

The resulting SDCC-FLP will have a triangular matrix of 12(12+1)/2 = 78 values, where only six of the random variables have cross correlation information with any of the other random variables. The size of the FLP can be reduced by using a Bit Vector to eliminate the uncorrelated values (note: the number of correlation values is 12(12-1)/2 = 66). Figure 20 illustrates the triangular matrix and the resulting SDCC-FLP that would be constructed.

In this matrix six of the standard deviation values are marked as not-measured (via method defined in Parent Document) and most of the cross correlation coefficient values are zero. The resulting SDCC-FLP contains the (see Table 4) Matrix Size (orange), Parse Control (grey) information and the Bit Vector (red); next the 12 standard deviation values (green) are included (with six of them marked as not measured); at the end of the SDCC-FLP six correlation coefficient values (blue) are included (all of the other correlation coefficients values are zero).

Table 5: Option 1 – Example Local Set: no Tag rearranging

Tag ID Number	Name	Standard Deviation and Correlation Coefficient Applicable (For Example Purposes)
1	Sensor ECEF Position Component X	YES
2	Sensor ECEF Position Component Y	YES
3	Sensor ECEF Position Component Z	YES
4	Sensor ECEF Velocity Component X	NO
5	Sensor ECEF Velocity Component Y	NO
6	Sensor ECEF Velocity Component Z	NO
7	Sensor Heading Angle	YES
8	Sensor Pitch Angle	YES
9	Sensor Roll Angle	YES
10	Sensor Absolute Heading Rate	NO
11	Sensor Absolute Pitch Rate	NO
12	Sensor Absolute Roll Rate	NO
ST 1010 SDCC-FLP N = 12		

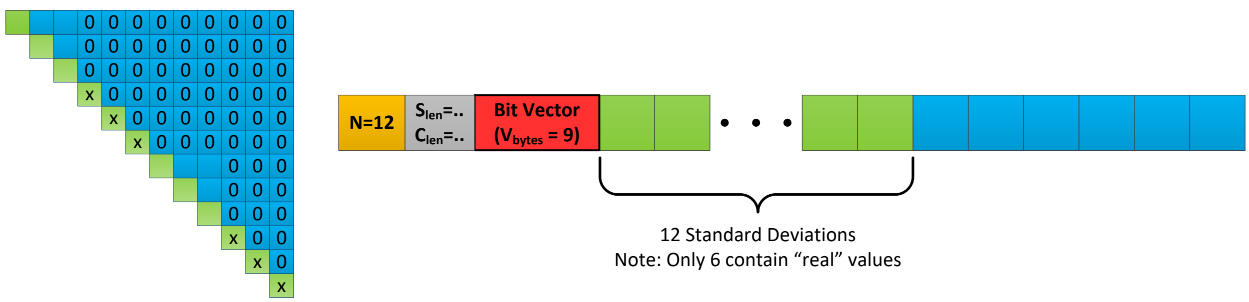

Figure 20: Illustration of Option 1 Example

8.3.1.2 Option 2

Option 2 rearranges the tag IDs, so the six random variables requiring standard deviations and correlation coefficients data are sent together in the KLV stream (see Table 6).

Table 6: Option 2 – Example Local Set: Tags arranged in two groups

Tag ID Number	Name	Standard Deviation and Correlation Coefficient Applicable (For Example Purposes)
1	Sensor ECEF Position Component X	YES
2	Sensor ECEF Position Component Y	YES
3	Sensor ECEF Position Component Z	YES
7	Sensor Heading Angle	YES
8	Sensor Pitch Angle	YES
9	Sensor Roll Angle	YES
ST 1010 SDCC-FLP N = 6		
4	Sensor ECEF Velocity Component X	NO
5	Sensor ECEF Velocity Component Y	NO
6	Sensor ECEF Velocity Component Z	NO

10	Sensor Absolute Heading Rate	NO
11	Sensor Absolute Pitch Rate	NO
12	Sensor Absolute Roll Rate	NO

The resulting SDCC-FLP will have a triangular matrix of 6(6+1)/2 = 21 values where the six values have cross correlation information in groups of three. The size of the FLP can be reduced by using a Bit Vector to eliminate the uncorrelated values. Figure 21 illustrates the triangular matrix and the resulting SDCC-FLP that would be constructed. In the matrix six standard deviations are represented; most of the cross correlation values are zero. The resulting SDCC-FLP contains the (see Table 4) Matrix Size (orange), Parse Control (grey) information and the Bit Vector (red); next the six Standard Deviation values (green) are included; at the end of the SDCC-FLP six Correlation Coefficients (blue) are included (all of the other Correlation Coefficients are zero).

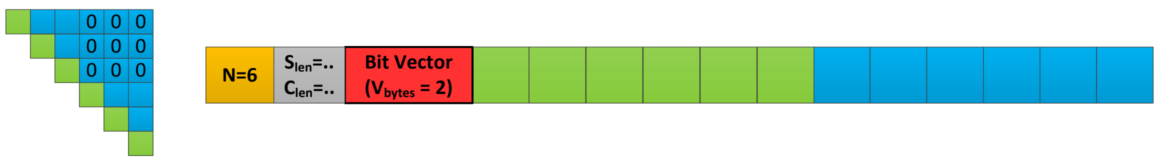

Figure 21: Illustration of Option 2 Example

8.3.1.3 Option 3

Option 3 has multiple SDCC-FLPs in the KLV stream. The 12 tag IDs may be strategically organized into multiple groups (see Table 7).

Table 7: Option 3 – Example Local Set: Tags arranged in three groups

Tag ID Number	Name	Standard Deviation and Correlation Coefficient Applicable (For Example Purposes)
1	Sensor ECEF Position Component X	YES
2	Sensor ECEF Position Component Y	YES
3	Sensor ECEF Position Component Z	YES
ST 1010 SDCC-FLP N = 3		
4	Sensor ECEF Velocity Component X	NO
5	Sensor ECEF Velocity Component Y	NO
6	Sensor ECEF Velocity Component Z	NO
7	Sensor Heading Angle	YES
8	Sensor Pitch Angle	YES
9	Sensor Roll Angle	YES
ST 1010 SDCC-FLP N = 3		
10	Sensor Absolute Heading Rate	NO
11	Sensor Absolute Pitch Rate	NO
12	Sensor Absolute Roll Rate	NO

The resulting Local Set will have two SDCC-FLPs each with a triangular matrix of 3(3+1)/2 = 6 values and all cross correlation information non-zero valued. Figure 22 illustrates the two triangular matrices and the resulting two SDCC-FLPs that would be constructed. Each resulting

SDCC-FLP contains the (see Table 4) Matrix Size (orange) and Parse Control (grey) information; next the three Standard Deviation values (green) are included; at the end of the SDCC-FLP three Correlation Coefficients (blue) are included.

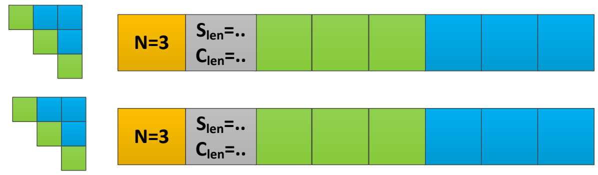

Figure 22: Illustration of Option 3 Example

In this example, two SDCC-FLPs are included in the Parent Local Set. This method also enables the two sets to have different types and sizes (i.e. C_f, C_{len}, S_f, S_{len}).

9 Deprecated Requirements

Requirement(s)	
ST 1010.1-03 (Deprecated)	A full diagonal of standard deviation values shall always be present in the SDCC-FLP.
ST 1010.1-07 (Deprecated)	The Standard Deviation and Correlation Coefficient FLP that describes a specific set of parameters in the invoking Standard or method shall follow those specific parameters in the invoking Standard or method in the KLV stream

Appendix A - Bit Vector Analysis

When using the Bit Vector there is the overhead cost of the Bit Vector itself. This analysis shows the tradeoff in using the Bit Vector to represent a sparse matrix versus this overhead cost.

The overhead in adding a Bit Vector is the number of bytes in the Bit Vector, which is a function of matrix size defined as V_{bytes} in Equation 5. The savings achieved by using the Bit Vector is computed by multiplying the number of zeros, Z, in the correlation coefficient matrix times the size, C_{len}, of each element. When this savings is greater than the overhead cost of the Bit Vector itself, then the Bit Vector is worth including. This relation is shown in Equation 6.

$$Z * C_{len} > V_{bytes} = \left\lceil \frac{M}{8} \right\rceil \geq \frac{M}{8} \qquad \textbf{Equation 6}$$

A more useful cost function (Equation 7) is found in rewriting Equation 6, wherein use of the Bit Vector is beneficial when the ratio of zeros (Z / M) multiplied by eight times the coefficient size (C_{len}) is greater than then one.

$$\frac{Z}{M} * (8C_{len}) > 1 \qquad \textbf{Equation 7}$$