H.264 / MPEG 4 AVC Decoder Implementation

Fanyu Ran[a], Yang Zhou[b], and Yifei Zhou[c]

[a]Student No.8657223, University of Ottawa
[b]Student No.8657223, University of Ottawa
[c]Student No.8635051, University of Ottawa

ABSTRACT

This document is p repared using LaTeX2e[1] and shows the desired format and appearance of a manuscript prepared for the Proceedings of the SPIE.[*] It contains general formatting instructions and hints about how to use LaTeX. The LaTeX source file that produced this document, `article.tex` (Version 3.4), provides a template, used in conjunction with `spie.cls` (Version 3.4). These files are available on the Internet at https://www.overleaf.com. The font used throughout is the LaTeX default font, Computer Modern Roman, which is equivalent to the Times Roman font available on many systems.

Keywords: Manuscript format, template, SPIE Proceedings, LaTeX

1. INTRODUCTION

H.264/AVC is the latest in a series of standards published by the ITU (International Telecommunications Union) and ISO (International Standards Organization). It describes and defines a method of coding video that can give better performance. H.264 makes it possible to compress video into a smaller space, which means that a compressed video clip takes up less transmission bandwidth and/or less storage space compared to older codecs. H.264 compression makes it possible to transmit HD television over a limited-capacity broadcast channel, to record hours of video on a Flash memory card and to deliver massive numbers of video streams over an already busy internet.

2. PYTHON PROGRAMMING LANGUAGE OVERVIEW

This section describes the normal structure of a manuscript and how each part should be handled. The appropriate vertical spacing between various parts of this document is achieved in LaTeX through the proper use of defined constructs, such as `\section{}`. In LaTeX, paragraphs are separated by blank lines in the source file.

At times it may be desired, for formatting reasons, to break a line without starting a new paragraph. This situation may occur, for example, when formatting the article title, author information, or section headings. Line breaks are inserted in LaTeX by entering `\\` or `\linebreak` in the LaTeX source file at the desired location.

2.1 Title and Author Information

The article title appears centered at the top of the first page. The title font is 16 point, bold. The rules for capitalizing the title are the same as for sentences; only the first word, proper nouns, and acronyms should be capitalized. Avoid using acronyms in the title. Keep in mind that people outside your area of expertise might read your article. At the first occurrence of an acronym, spell it out, followed by the acronym in parentheses, e.g., noise power spectrum (NPS).

The author list is in 12-pt. regular, centered. Omit titles and degrees such as Dr., Prof., Ph.D., etc. The list of affiliations follows the author list. Each author's affiliation should be clearly noted. Superscripts may be used to identify the correspondence between the authors and their respective affiliations. Further author information, such as e-mail address, complete postal address, and web-site location, may be provided in a footnote by using `\authorinfo{}`, as demonstrated above.

Further author information:
Fanyu Ran: E-mail: aaa@tbk2.edu
Yang Zhou: E-mail: bba@cmp.com
Yifei Zhou: E-mail: yzhou151@uottawa.ca

[*]The basic format was developed in 1995 by Rick Hermann (SPIE) and Ken Hanson (Los Alamos National Lab.).

2.2 Abstract and Keywords

The title and author information is immediately followed by the Abstract. The Abstract should concisely summarize the key findings of the paper. It should consist of a single paragraph containing no more than 250 words. The Abstract does not have a section number. A list of up to eight keywords should immediately follow the Abstract after a blank line. These keywords will be included in a searchable database at SPIE.

2.3 Body of Paper

The body of the paper consists of numbered sections that present the main findings. These sections should be organized to best present the material. See Sec. **??** for formatting instructions.

2.4 Appendices

Auxiliary material that is best left out of the main body of the paper, for example, derivations of equations, proofs of theorems, and details of algorithms, may be included in appendices. Appendices are enumerated with uppercase Latin letters in alphabetic order, and appear just before the Acknowledgments and References. Appendix A contains more about formatting equations and theorems.

2.5 Acknowledgments

In the Acknowledgments section, appearing just before the References, the authors may credit others for their guidance or help. Also, funding sources may be stated. The Acknowledgments section does not have a section number.

2.6 References

SPIE is able to display the references section of your paper in the SPIE Digital Library, complete with links to referenced journal articles, proceedings papers, and books, when available. This added feature will bring more readers to your paper and improve the usefulness of the SPIE Digital Library for all researchers. The References section does not have a section number. The references are numbered in the order in which they are cited. Examples of the format to be followed are given at the end of this document.

The reference list at the end of this document is created using BibTeX, which looks through the file `report.bib` for the entries cited in the LaTeX source file. The format of the reference list is determined by the bibliography style file `spiebib.bst`, as specified in the `\bibliographystyle{spiebib}` command. Alternatively, the references may be directly formatted in the LaTeX source file.

For books,[1–3] the listing includes the list of authors, book title, publisher, city, page or chapter numbers, and year of publication. A reference to a journal article[4] includes the author list, title of the article (in quotes), journal name (in italics, properly abbreviated), volume number (in bold), inclusive page numbers, and year. By convention,[1] article titles are capitalized as described in Sec. 2.1. A reference to a proceedings paper or a chapter in an edited book[5] includes the author list, title of the article (in quotes), volume or series title (in italics), volume number (in bold), if applicable, inclusive page numbers, publisher, city, and year. References to an article in the SPIE Proceedings may include the conference name (in italics), as shown in Ref. 6. For websites[7] the listing includes the list of authors, title of the article (in quotes), website name, article date, website address either enclosed in chevron symbols ('<' and '>'), underlined or linked, and the date the website was accessed.

If you use this formatting, your references will link your manuscript to other research papers that are in the CrossRef system. Exact punctuation is required for the automated linking to be successful.

Citations to the references are made using superscript numerals, as demonstrated in the above paragraph. One may also directly refer to a reference within the text, e.g., "as shown in Ref. 4 ..."

2.7 Footnotes

Footnotes[†] may be used to provide auxiliary information that doesn't need to appear in the text, e.g., to explain measurement units. They should be used sparingly, however.

Only nine footnote symbols are available in LaTeX. If you have more than nine footnotes, you will need to restart the sequence using the command `\footnote[1]{Your footnote text goes here.}`. If you don't, LaTeX will provide the error message `Counter too large.`, followed by the offending footnote command.

3. H.264 DECODING OVERVIEW

3.1 H.264 Profiles

H.264 specification includes several profiles, each specifying a subset of the optional features in the H.264 standard. A Profile places limits on the algorithmic capabilities required of an H.264 decoder. Each Profile is intended to be useful to a class of applications. For example, the Baseline Profile may be useful for low-delay, such as video conferencing, with relatively low computational requirements. The Main Profile may be suitable for basic television/entertainment applications such as Standard Definition TV services. The High Profiles add tools to the Main Profile which can improve compression efficiency especially for higher spatial resolution services, such as High Definition TV.

3.2 Main Decoding Processes

The decoding process is carried out by an H.264 video decoder. Flow diagram is shown in Figure 1. (1) Initially, decoder codes compressed bitstream and does entropy decoding, obtaining a coefficient matrix. The following steps are dequantization and IDCT. Residual is output of this stage. Then the decoder will decode the frame depending on the combination of residual and prediction results. Notice that, in the both of coder and decoder, data is processed in units of macroblock. A decoded residual macroblock is formed by re-scaling and inverse transformation. Prediction is created in decoder and is added to the residual to generated a decoded macroblock.

3.2.1 Parsing Process with Entropy Decoding

Inputs to parsing process are bitstream from contents of video slices. Output of this process are syntax element values. (Video sequence that is represented in a particular format which is synatx). (2) Actually, several methods are specified by H.264 for coding symbols, including Fixed Length Code, Exponential-Golomb variable length code, CAVLC and CABAC.

And the data can be divided into two parts: decoder controlling parameters and transform coefficients. Above the slice data level, bits are decoded using Fixed Length decoder or Exp-Golomb decoder. In the slice data level and below, coefficient values are decoded in one of two ways. In CABAC mode, all symbols are decoded by CABAC decoder. Otherwise, coefficient values using CAVLC and other symbols are coded using fixed length or Exp-Golomb decoder. In our baseline decoder, Exp-Golomb and CAVLC are utilized.

3.2.2 Exponential Golomb Decoder

Exponential Golomb codes (ExpG) are an efficient way of representing data symbols with varying probabilities. It assigns short codewords to frequently-occurring data symbols and long codewords to less common data symbols. A parameter code_num indicates codewords. The length of Exponential-Golomb codes are varying with two properties:

1. With the increase of code_num, code length also rises.
2. Without looking up for tables, each code can be generated and decode algorithmically.

[†] Footnotes are indicated as superscript symbols to avoid confusion with citations.

3.2.3 Context Adaptive Variable Length Coding (CAVLC)

Residual, the transform coefficients are decoded by CAVLC decoder. Each 8×8 block of quantized transform coefficients is processed as four 4×4 blocks for the purposes of CAVLC encoding and decoding. The decoding process can be divided into 5 stages:

1. Decode the number of coefficients and trailing ones
2. Decode the sign of each trailing ones
3. Decode the levels of the remaining non-zero coefficients
4. Decode the number of zeros preceding each non-zero coefficient
5. Zero padding

Decode the number of coefficients and trailing ones, according to input bitstream and standard table 9-5. Then the result is used to decode sign of trailing ones. To decode the levels of the remaining coefficients, we need to obtain the suffix length and get level prefix depending on standard table 9-6. The next step is to decode the number of zeros preceding each non-zero coefficient. The total number of coefficients and standard table 9-7 are needed. After the four steps, use zero to pad the remaining elements.

3.3 Prediction Overview

The prediction methods may have a great influence on the compression performance.H.264 supports two prediction options: **Intra prediction** using data within the current frame, **Inter prediction** using motion compensated prediction from previously coded frames. H.264 provides multiple prediction block sizes, multiple reference frames and special modes. All these features give H.264 a great deal of flexibility in the prediction process. By selecting the best prediction options for an individual macroblock, H.264 can minimize the residual size to produce a highly compressed bitstream.

3.3.1 Macroblocks In Prediction

H.264 divides each frame in the video into macroblocks, which are blocks of 16x16 pixels. The macroblocks are decoded from left to right, and then top to bottom. There are three types of marcoblocks for prediction sources, an I Macroblock, a P Macroblock and a B Macroblock. An I Macroblock (I MB) is predicted using intra prediction from neighbouring samples in the current frame. A P Macroblock (P MB) is predicted from samples in a previously-coded frame which may be before or after the current picture in display order. Each partition in a B Macroblock (B MB) is predicted from samples in one or two previously-coded frames.

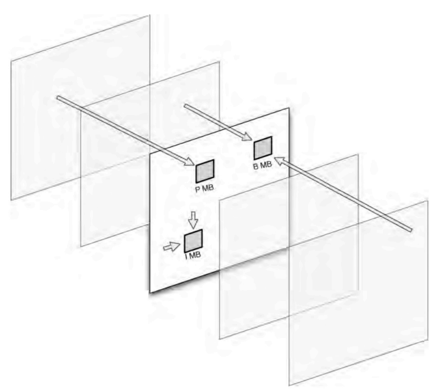

Figure 1. Marcoblock types

3.3.2 Intra-Prediction

Intra prediction is used to use part of the frame to predict the other parts. An I Macroblock (I MB) is predicted using intra prediction from neighbouring samples in the current frame. For every block of the frame up to 16x16 pixels, intraprediction uses previously decoded neighbor blocks to give an estimate for the new block.

In an intra macroblock, there are three choices of intra prediction block size for the luma component, namely 16×16, 8×8 or 4×4. A single prediction block is generated for each chroma component. Each prediction block is generated using one of a number of possible prediction modes.

Intra prediction block size	Notes
16×16 (luma)	A single 16×16 prediction block P is generated.
8×8 (luma)	An 8×8 prediction block P is generated for each 8×8 luma block.
4×4 (luma)	A 4×4 prediction block P is generated for each 4×4 luma block.
Chroma	One prediction block P is generated for each chroma component. Four possible prediction mod

Table 1. Intra prediction types

4×4 luma has 9 prediction modes,The samples, labelled A-M, have previously been decoded and reconstructed and they act as the prediction reference, then decoder copies the reference samples to the currently decoding blocks to continue the decode process. and decoder to form a prediction reference.

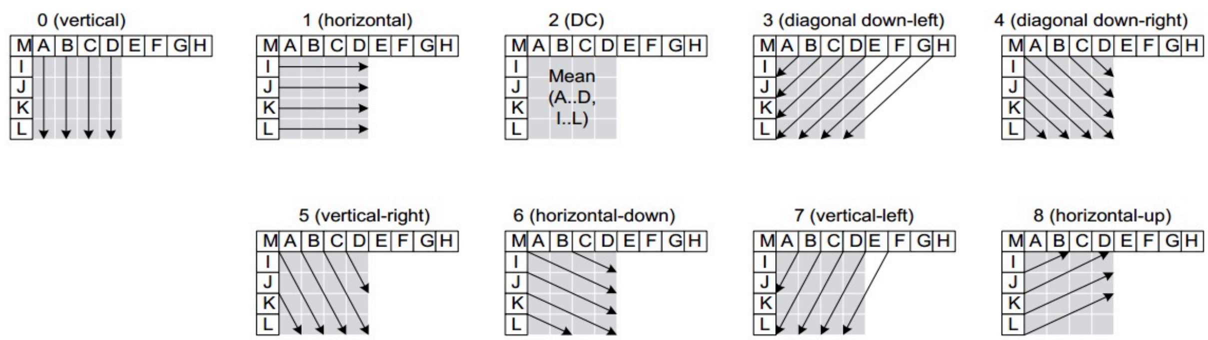

Figure 2. 4×4 intra prediction modes

As an alternative to the 4 4 luma modes, the entire 16 16 luma component of a macroblock may be predicted in one operation. Four modes are available.

Intra prediction of the luma component with an 8×8 block size is only available in the High profiles. Each 8×8 luma block in a macroblock is predicted using one of nine prediction modes which are as same as the nine modes of 4×4 luma.

Each chroma component of a macroblock is predicted from previously encoded chroma samples above and/or to the left, with both chroma components always using the same prediction mode. The four prediction modes are similar to the 16 16 luma prediction modes. The numbering of the modes is different, but the rules are in common.

3.3.3 Inter-Prediction

Inter prediction is the process of predicting a block of luma and chroma samples from a reference picture that has previously been coded and transmitted. It takes advantage of the fact that the content of a new frame in the video often has high correlation to the data in the previous frames.The offset between the position of the

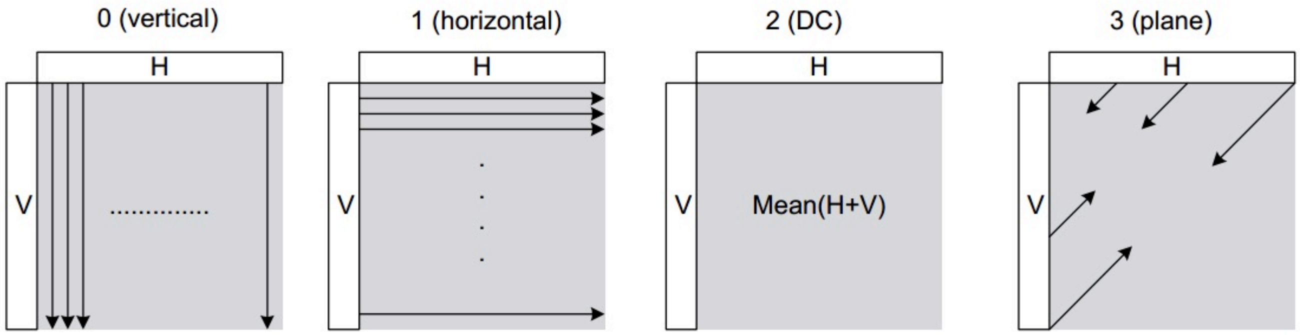

Figure 3. 4×4 intra prediction modes

current partition and the prediction region in the reference picture is a motion vector. The motion vector may point to integer, half- or quarter-sample positions in the luma component of the reference picture.

The decoded pictures stored in the Decoded Picture Buffer (DPB), in which case they may be used as reference pictures for inter prediction.The pictures in the DPB are listed in a particular order, and the list can be classified into three different types:

List0 (P slice)	A single list of all the reference pictures. By default, the first picture in the List is the most recently decod
List0 (B slice)	A list of all the reference pictures. By default, the first picture in the List is the picture **before** the curren
List1 (B slice)	A list of all the reference pictures. By default, the first picture in the List is the picture **after** the current

Table 2. Three types of DPB list

Each 16×16 P or B macroblock may be predicted using a range of block sizes. The macroblock is split into one, two or four macroblock partitions:(a) one 16×16 macroblock partition (b) two 8×16 macroblock partitions (c) two 16×8 macroblock partitions (d) four 8×8 macroblock partitions

If 8×8 partition size is chosen, then each 8×8 block of luma samples and associated chroma samples, a sub-macroblock, is split into one, two or four sub-macroblock partitions): one 8×8, two 4×8, two 8×4 or four 4×4 sub-MB partitions.

Each macroblock partition and sub-macroblock partition has one or two motion vectors (x, y). Each of them points to an area of the same size in a reference frame that is used to predict the current partition. Motion vectors for neighboring partitions are often highly correlated and so each motion vector is predicted from vectors of nearby, previously coded partitions. Motion vector is supposed to compensate the decoded picture from the reference ones. It means we could transmit the motion vector instead of the whole redundant of reference marcoblock.

4. SOFTWARE DESIGN OVERVIEW

Figures are numbered in the order of their first citation. They should appear in numerical order and on or after the same page as their first reference in the text. Alternatively, all figures may be placed at the end of the manuscript, that is, after the Reference section. It is preferable to have figures appear at the top or bottom of the page. Figures, along with their captions, should be separated from the main text by at least 0.2 in. or 5 mm.

Figure captions are centered below the figure or graph. Figure captions start with the figure number in 9-point bold font, followed by a period; the text is in 9-point normal font; for example, "Figure 3. Original image...". See Fig. 5 for an example of a figure caption. When the caption is too long to fit on one line, it should be justified to the right and left margins of the body of the text.

Tables are handled identically to figures, except that their captions appear above the table.

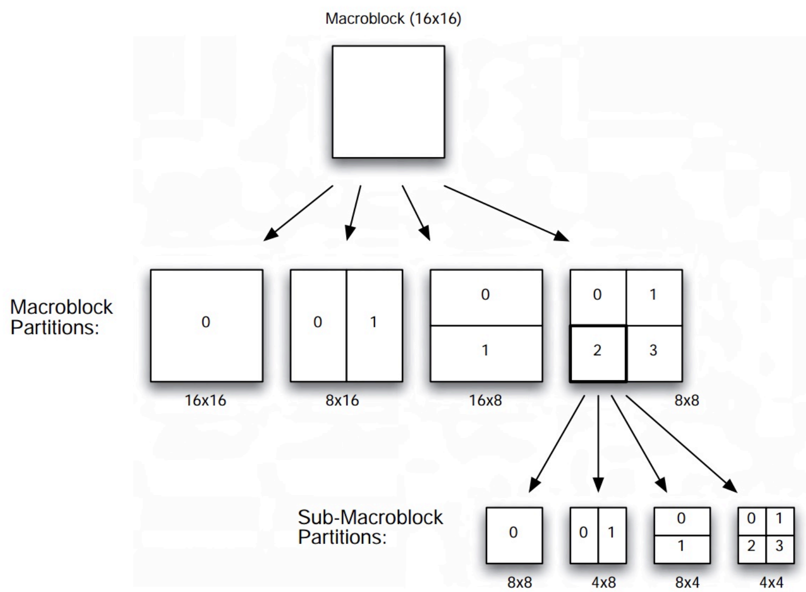

Figure 4. Macroblock partitions and sub-macroblock partitions

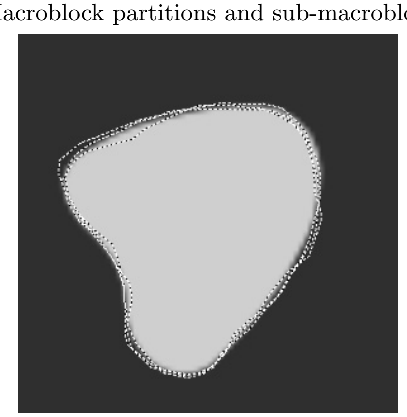

Figure 5. Figure captions are used to describe the figure and help the reader understand it's significance. The caption should be centered underneath the figure and set in 9-point font. It is preferable for figures and tables to be placed at the top or bottom of the page. LaTeX tends to adhere to this standard.

5. H.264 AND IMPLEMENTATION DETAILS

Video and audio files can be included for publication. See Tab. **??** for the specifications for the mulitimedia files. Use a screenshot or another .jpg illustration for placement in the text. Use the file name to begin the caption. The text of the caption must end with the text "http://dx.doi.org/doi.number.goes.here" which tells the SPIE editor where to insert the hyperlink in the digital version of the manuscript.

Here is a sample illustration and caption for a multimedia file:

6. RESULT AND DESIGN EXPLORATION

Begin the Introduction below the Keywords. The manuscript should not have headers, footers, or page numbers. It should be in a one-column format. References are often noted in the text and cited at the end of the paper.

LaTeX margins are related to the document's paper size. The paper size is by default set to USA letter paper. To format a document for A4 paper, the first line of this LaTeX source file should be changed to `\documentclass[a4paper]{spie}`.

Table 3. Fonts sizes to be used for various parts of the manuscript. Table captions should be centered above the table. When the caption is too long to fit on one line, it should be justified to the right and left margins of the body of the text.

Article title	16 pt., bold, centered
Author names and affiliations	12 pt., normal, centered
Keywords	10 pt., normal, left justified
Abstract Title	11 pt., bold, centered
Abstract body text	10 pt., normal, justified
Section heading	11 pt., bold, centered (all caps)
Subsection heading	11 pt., bold, left justified
Sub-subsection heading	10 pt., bold, left justified
Normal text	10 pt., normal, justified
Figure and table captions	9 pt., normal
Footnote	9 pt., normal
Reference Heading	11 pt., bold, centered
Reference Listing	10 pt., normal, justified

Table 4. Margins and print area specifications.

Margin	A4	Letter
Top margin	2.54 cm	1.0 in.
Bottom margin	4.94 cm	1.25 in.
Left, right margin	1.925 cm	.875 in.
Printable area	17.15 x 22.23 cm	6.75 x 8.75 in.

Authors are encouraged to follow the principles of sound technical writing, as described in Refs. 2 and 8, for example. Many aspects of technical writing are addressed in the *AIP Style Manual*, published by the American Institute of Physics. It is available on line at `https://publishing.aip.org/authors`. A spelling checker is helpful for finding misspelled words.

An author may use this LaTeX source file as a template by substituting his/her own text in each field. This document is not meant to be a complete guide on how to use LaTeX. For that, please see the list of references at `http://latex-project.org/guides/` and for an online introduction to LaTeX please see 7.

APPENDIX A. MISCELLANEOUS FORMATTING DETAILS

It is often useful to refer back (or forward) to other sections in the article. Such references are made by section number. When a section reference starts a sentence, Section is spelled out; otherwise use its abbreviation, for example, "In Sec. 2 we showed..." or "Section 2.1 contained a description...". References to figures, tables, and theorems are handled the same way.

A.1 Formatting Equations

Equations may appear in line with the text, if they are simple, short, and not of major importance; e.g., $\beta = b/r$. Important equations appear on their own line. Such equations are centered. For example, "The expression for the field of view is

$$2a = \frac{(b+1)}{3c}\,, \tag{1}$$

where a is the ..." Principal equations are numbered, with the equation number placed within parentheses and right justified.

Equations are considered to be part of a sentence and should be punctuated accordingly. In the above example, a comma follows the equation because the next line is a subordinate clause. If the equation ends the sentence, a period should follow the equation. The line following an equation should not be indented unless it is meant to start a new paragraph. Indentation after an equation is avoided in LaTeX by not leaving a blank line between the equation and the subsequent text.

References to equations include the equation number in parentheses, for example, "Equation (1) shows ..." or "Combining Eqs. (2) and (3), we obtain..." Using a tilde in the LaTeX source file between two characters avoids unwanted line breaks.

A.2 Formatting Theorems

To include theorems in a formal way, the theorem identification should appear in a 10-point, bold font, left justified and followed by a period. The text of the theorem continues on the same line in normal, 10-point font. For example,

Theorem 1. For any unbiased estimator...

Formal statements of lemmas and algorithms receive a similar treatment.

ACKNOWLEDGMENTS

First we would like to thank Professor Zhao for supervising our project. And then we would like to thank our classmates that had helped us during our project period.

REFERENCES

[1] Lamport, L., [*LaTeX: A Document Preparation System*], Addison-Wesley, Reading, Mass. (1994).

[2] Alred, G. J., Brusaw, C. T., and Oliu, W. E., [*Handbook of Technical Writing*], St. Martin's, New York (2015 (eleventh edition)).

[3] Goossens, M., Mittelbach, F., and Rahtz, S., [*The LaTeX Companion*], Addison-Wesley, Reading, Mass. (1997).

[4] Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., "Equations of state calculations by fast computing machine," *J. Chem. Phys.* **21**, 1087–1091 (1953).

[5] Gull, S. F., "Developments in maximum-entropy data analysis," in [*Maximum Entropy and Bayesian Methods*], Skilling, J., ed., 53–71, Kluwer Academic, Dordrecht (1989).

[6] Hanson, K. M., "Introduction to Bayesian image analysis," in [*Medical Imaging: Image Processing*], Loew, M. H., ed., *Proc. SPIE* **1898**, 716–731 (1993).

[7] Lees-Miller, J. D., "Free and Interactive Online Introduction to LaTeX." Overleaf, 26 February 2015 https://www.overleaf.com/latex/learn/free-online-introduction-to-latex-part-1. (Accessed: 15 September 2015).

[8] Perelman, L. C., Paradis, J., and Barrett, E., [*Mayfield Handbook of Technical and Scientific Writing*], Mountain View, Mayfield (April 1997). http://mit.imoat.net/handbook/.