UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

SAMSUNG ELECTRONICS CO., LTD.
Petitioner

v.

PROMOS TECHNOLOGIES, INC.
Patent Owner

Patent No. 6,163,492

DECLARATION OF INGRID HSIEH-YEE, PH.D. IN SUPPORT OF PETITION FOR *INTER PARTES* REVIEW OF U.S. PATENT NO. 6,163,492

Table of Contents

I, Ingrid Hsieh-Yee, Ph.D., declare as follows:

I. INTRODUCTION

1. I have been retained as an independent expert witness on behalf of Samsung Electronics Co., Ltd. ("Samsung") in this proceeding before the United States Patent and Trademark Office ("PTO") regarding U.S. Patent No. 6,163,492.

2. I am being compensated for my work in this matter at my accustomed hourly rate. I am also being reimbursed for reasonable and customary expenses associated with my work and testimony in this investigation. My compensation is not contingent on the results of my study, the substance of my opinions, or the outcome of this matter.

3. In the preparation of this declaration, I have reviewed the documents referenced below. Each of these documents is a type of material that experts in my field (as I discuss below) would reasonably rely upon when forming their opinions. I refer to these documents by exhibit numbers that I understand are assigned to them in this proceeding before the PTO.

(1) Weste, N. H. E. & Eshraghian, K. (1993). PRINCIPLES OF CMOS VLSI DESIGN : A SYSTEMS PERSPECTIVE ("Weste"), **Ex. 1014**;

(2) Bibliographic record for PRINCIPLES OF CMOS VLSI DESIGN : A SYSTEMS PERSPECTIVE (1993) available at the Library of Congress online catalog at https://lccn.loc.gov/92016564, accessed on February 26,

2019, **Appendix C**;

(3) MARC record for PRINCIPLES OF CMOS VLSI DESIGN : A SYSTEMS PERSPECTIVE (1993) available at the Library of Congress online catalog at https://catalog.loc.gov/vwebv/staffView?searchId=23680&recPointer=0&recCount=25&searchType=1&bibId=1559597, accessed on February 26, 2019, **Appendix D**;

(4) Copyright registration record for PRINCIPLES OF CMOS VLSI DESIGN : A SYSTEMS PERSPECTIVE (1993), available at the public catalog of the United States Copyright Office at https://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?v1=1&ti=1,1&Search%5FArg=principles%20of%20cmos%20vlsi%20design&Search%5FCode=TALL&CNT=25&PID=Pgd2v2jBNnZG4h6jx7oX1FhZaj&SEQ=20190303114850&SID=3, accessed on February 26, 2019, **Appendix E**.

4. In forming the opinions expressed within this declaration, I have considered:

(1) The documents listed above;

(2) The reference materials cited herein; and

(3) My own academic background and professional experiences, as described below.

5. My complete qualifications and professional experience are described in my academic curriculum vitae, a copy of which is attached as **Appendix A**. The following is a brief summary of my relevant qualifications and professional

experience.

6. I am currently a Professor in the Department of Library and Information Science at the Catholic University of America. I have experience working in an academic library, a medical library, and a legislative library and have been a professor for more than 25 years. I hold a Ph.D. in Library and Information Studies from the University of Wisconsin-Madison and a Masters degree in Library and Information Studies from the University of Wisconsin-Madison.

7. I am an expert on library cataloging and classification and have published two books on this subject, *Organizing Audiovisual and Electronic Resources for Access*: *A Cataloging Guide* (2000, 2006). I teach a variety of courses, including Cataloging and Classification, Internet Searches and Web Design, Advanced Cataloging and Classification, Organization of Internet Resources, Advanced Information Retrieval and Analysis Strategies, the Information Professions in Society, Information Literacy Instruction, and Digital Content Creation and Management. My research interests cover cataloging and classification, information organization, metadata, information retrieval, information architecture, digital collections, scholarly communication, user interaction with information systems, and others.

8. I am fully familiar with a library cataloging encoding standard known as the "Machine-Readable Cataloging" standard, also known as "MARC," which became the national standard for sharing bibliographic data in the United States by 1971 and the international standard by 1973. MARC is the primary communications protocol for the transfer and storage of bibliographic metadata in libraries. Experts in my field would reasonably rely upon MARC records when forming their opinions.

9. A MARC record comprises several fields, each of which contains specific data about the work. Each field is identified by a standardized, unique, three-digit code corresponding to the type of data that follows. **Appendix B** is a true and correct copy of Parts 7 to 10 of "Understanding MARC Bibliographic: Machine-Readable Cataloging" (http://www.loc.gov/marc/umb/) from the Library of Congress that explains commonly used MARC fields. For example, the personal author of the work is recorded in Field 100, the title is recorded in Field 245, publisher information is recorded in Field 260, and the physical volume and characteristics of a publication are recorded in Field 300, and topical subjects are recorded in the 650 fields.

10. Library online catalogs are based on MARC records that represent their collections and help the public understand what materials are publicly

accessible in those libraries. Most libraries with online catalogs have made their catalogs freely available on the Web. These online catalogs offer user-friendly search interfaces, often in the form of a single search box, to support searching by author, title, subject, keywords and other data elements. They also offer features for users to narrow their search results by language, year, format, and other elements. Many libraries display MARC records on their online catalogs with labels for the data elements to help the public interpret MARC records. Many libraries also offer an option to display MARC records in MARC fields. For monographs, after a MARC record is created and made searchable on a library catalog, it is customary library practice to have the physical volume processed for public access soon after, usually within a week.

11. I used authoritative information systems such as the online catalog at the Library of Congress (https://catalog.loc.gov) and the public catalog of the United States Copyright Office (http://cocatalog.loc.gov/cgi-bin/Pwebrecon.cgi?DB=local&PAGE=First) to search for records. These records are identified and discussed in this declaration. Experts in the field would reasonably rely on the data described herein to form their opinions.

II. EX. 1014 WESTE

12. **Ex. 1014** is a true and correct copy of portions of Weste. When I was asked to prepare this declaration, I searched the Library of Congress online catalog for records because books published in the United States tend to be registered with the Copyright Office of the Library of Congress and cataloged by the Library of Congress. The search results informed me that the Library of Congress provided access to Weste. The "Item Availability" area of the library records indicates that the Library of Congress holds one copy of this book, which is "stored offsite" and can be requested in the "Jefferson or Adams Building Reading Rooms." The record I saved on February 26 shows the item status was "c.1 Charged – Due – (Internal Loan)," meaning it was being used at the time. I requested access to this title later and received the physical volume of copy 1 at the Science and Technology Reading Rooms in the Library of Congress Adams Building on March 1, 2019.

13. **Ex. 1014** is a true and correct copy of portions of Weste that I made on March 1, 2019, while the book was in my possession at the Library of Congress. I obtained **Ex. 1014** by personally scanning the cover, series title page, title page, copyright page, about the author page, preface, table of contents, the last numbered page of the book, back cover, and spine of this book. These pages are presented as **Ex. 1014**.

14. The cover of **Ex. 1014** shows that "PRINCIPLES OF CMOS VLSI

DESIGN : A Systems Perspective" is the title, "Neil H. E. Weste" and "Kamran Eshraghian" are the authors, and the publication is the "second edition." The cover also shows a Library of Congress bar code, with the number "00001496268." The series title page of **Ex. 1014** shows that the book is in the "Addison-Wesley VLSI Systems Series." The title page of **Ex. 1014** shows the same title, authors, and edition number as the information presented on the cover. It also shows that "Addison-Wesley Publishing Company" of Reading, Massachusetts is the publisher. The copyright page of **Ex. 1014** carries a stamp of "LIBRARY OF CONGRESS MAY 13 1993 CIP," where CIP stands for Cataloging in Publication. It also shows a block of "Library of Congress Cataloging-in-Publication Data," a hand-written call number of "TK7874 .W46 1993," a copyright date of "1993" held by AT&T. The back cover shows the title's ISBN (International Standard Book Number) is "0-201-53376-6" and the spine of the book has a label of the call number "TK 7874 .W46 1993" and a label of its location "CBN BRNCH GenColl," meaning Cabin Branch general collection.

A. Library of Congress – Bibliographic and MARC Records

15. **Appendix C** is a true and correct copy of the Bibliographic record for Weste. I retrieved this record from the Library of Congress online catalog (https://catlog.loc.gov) after searching for the book by its title, "principles of CMOS VLSI design." I personally identified and located this record, which experts

in my field would reasonably rely upon when forming their opinions.

16. **Appendix D** is a true and correct copy of the MARC record for Weste. I retrieved this record from the Library of Congress online catalog (https://catlog.loc.gov) after searching for the book by its title, "principles of CMOS VLSI design" and selecting the MARC Tags display. I personally identified and located this record, which experts in my field would reasonably rely upon when forming their opinions.

17. The main title field of the Bibliographic record (**Appendix C**) and Field 245 of the MARC record (**Appendix D**) identify the book title as "Principles of CMOS VLSI design : a systems perspective" and identify "Neil H. E. Weste, Kamran Eshraghian" as the authors. The edition area of the Bibliographic record and Field 250 of the MARC record identify this publication as the "2nd ed." The published/created field of the Bibliographic record and Field 260 of the MARC record show that Weste was published by Addison-Wesley Pub. Co. of Reading, Mass. in 1993. The ISBN field of the Bibliographic record and Field 020 of the MARC record show a number of "0201533766," which matches the ISBN on the back cover of **Ex. 1014**. The description field of the Bibliographic record and Field 300 of the MARC record show the book has "xxii" introductory pages, and the last numbered page is page 713, which matches the last numbered page of **Ex. 1014**.

The title, authors, edition number, publisher, publication date, ISBN and pagination recorded in the Bibliographic and MARC records match the information contained in **Ex. 1014**. Field 991 of the MARC record shows a code of "00001496268," which matches the label on the cover of **Ex. 1014**. It also shows the book is in the "GenColl" with a number of "TK7874 .W46 1993," which matches the spine label of **Ex. 1014**.

18. The first six digits of Field 008 of the MARC record (**Appendix D**) inform me that the record for the book was first created on "920421" (*i.e.*, April 21, 1992) and Field 955, a field for recording local cataloging dates, shows "CIP ver. pv07 06-16-93" (*i.e.*, June 16, 1993). Field 040 subfield "a" shows that "DLC" is the library that created the original MARC record. According to the Directory of OCLC Members (https://www.oclc.org/en/contacts/libraries.html), "DLC" is the OCLC symbol for the Library of Congress. These data inform my opinion that the MARC record (**Appendix D**) was first created on April 21, 1992, as a Cataloging-in-Publication record based on the information presented by the publisher. After the CIP record was completed, it was sent to the publisher for inclusion in the published volume. The block of "Library of Congress Cataloging-in-Publication Data" on the copyright page of **Ex. 1014** reflects this practice. That block of data includes a number of "92-16564," which has the appearance of a Library of Congress control number. This number is reflected in the LCCN field of the

Bibliographic record and Field 010 of the MARC record. Following the Cataloging in Publication (CIP) Program agreement, the publisher sent the physical volume to the CIP after the volume was published. CIP catalogers then used the volume to verify the record on June 16, 1993.

19. Field 050 of the MARC record (**Appendix D**) shows Weste has been assigned a Library of Congress Classification (LCC) number, "TK7874," which is the class number for "Microelectronics. Integrated circuits." Field 082 shows the Dewey Decimal Classification (DDC) Number assigned to this publication is "621.3/95," which is the class number for "circuity." The subjects of this book are also represented by Library of Congress subject headings in two Field 650s: "Integrated circuits |x Very large scale integration |x design and construction" and "Metal oxide semiconductors, Complementary" (|x represents a topical subdivision). Because the book is in a series, the series title, "The VLSI systems series," is also recorded in Field 440 to make the series title searchable. Field 700 records the authorized form of the second author and makes this author an access point for this book.

20. This MARC record (**Appendix D**) has made Weste (**Ex. 1014**) searchable in the Library of Congress online catalog. As a result, users interested in topics covered by this book can find it by the LCC number, the DDC number, the

subject headings and the series title. In addition, users can find this book by its title, authors, and ISBN.

21. Based on the information above, it is my opinion that PRINCIPLES OF CMOS VLSI DESIGN : A SYSTEMS PERSPECTIVE (**EX. 1014**) is a book that has been made available by the Library of Congress, meaning that anyone who was interested in the topics covered by the book would have been able to search for and access PRINCIPLES OF CMOS VLSI DESIGN : A SYSTEMS PERSPECTIVE.

B. <u>Library of Congress – Date Stamp</u>

22. The copyright page of Weste in **Ex. 1014** bears a stamp of "LIBRARY OF CONGRESS MAY 13 1993 CIP" that indicates the date when the Cataloging in Publication Program of the Library of Congress received the physical volume of Weste from the publisher. The stamp has the appearance and distinctive characteristics of a typical check-in date stamp utilized by libraries to indicate the date a particular publication was received by the library. The date stamp and the cataloging dates in Field 955 of the MARC record (**Appendix D**) inform my opinion that the MARC record was created by Library of Congress catalogers on "920421" (*i.e.*, April 21, 1992) as a CIP record, that the physical volume was received by the CIP Program of the Library of Congress on May 13, 1993, and that CIP catalogers used the physical volume to verify the record on "06-16-1993" (*i.e.*, June 16, 1993). In most academic libraries a newly cataloged book

would have become available for the public soon after the cataloging record is completed, usually within a week. Considering the volume of materials the Library of Congress needs to catalog and process, my conservative estimate is that Weste would have become available for public access by September 16, 1993, three months after the cataloging record was verified.

23. Based on the date stamp placed on the copyright page in Weste (**Ex. 1014**), which has a date stamp of May 13, 1993, the Bibliographic record (**Appendix C**), the MARC record (**Appendix D**), and my understanding of the ordinary and customary cataloging and processing practices of libraries, it is my opinion that Weste (**Ex. 1014**) was cataloged and became searchable as early as April 1992 when the CIP record was created, and the physical item would have become accessible to the public by September 1993, three months after the cataloging record was verified in June 1993.

C. United States Copyright – Copyright Registration

24. **Appendix E** is a true and accurate copy of the copyright registration record obtained from the public catalog of the United States Copyright Office in the Library of Congress. I obtained **Appendix E** by first accessing the http://www.copyright.gov/ website, selecting the hyperlink "Search Copyright Records," and performing a search by the title on February 26, 2019, to retrieve the record for "Principles of CMOS VLSI design." I personally identified and located

this record, which experts in my field would reasonably rely upon when forming their opinions.

25. The copyright record shows that PRINCIPLES OF CMOS VLSI DESIGN : A SYSTEMS PERSPECTIVE is the title of this book and that "Neil H. E. Weste, Kamran Eshraghian" are the authors. The record identifies the book as the "2nd ed." It also shows Addison-Wesley Pub. Co., of Reading, MA as the publisher. It shows the book has a copyright date of 1993 and "AT&T Bell Telephone Laboratories, Inc." as the copyright claimant. The date of creation is "1992," the date of publication is "1993-04-26" (*i.e.*, April 26, 1993) and the copyright registration date is "1993-08-20" (*i.e.*, August 20, 1993). The title, authors, edition number, publisher, publication date and the copyright holder information on the copyright record match the information contained in **Ex. 1014**. The copyright record also shows that Weste has 713 pages, which matches the last numbered page of this book recorded in Field 300 of the MARC record (**Appendix D**) and is consistent with the page number of the last page of Weste (**Ex. 1014**), which I personally scanned at the Library of Congress.

D. Actual Usage

26. Actual usage of a publication is reflected by the papers that make reference to it. My research on Google Scholar has identified at least three articles published from 1994 to 1996 that cited PRINCIPLES OF CMOS VLSI DESIGN : A

SYSTEMS PERSPECTIVE (1993). The earliest citing document was published in December 1994. The citations of the citing articles are listed here for easy reference, and these citing articles are attached in **Appendix F**.

(1) Cong, J., & Koh, C. K. (1994). Simultaneous driver and wire sizing for performance and power optimization. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, *2*(4), 408-425.

(2) Kong, J. T., & Overhauser, D. (1995). Methods to improve digital MOS macromodel accuracy. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, *14*(7), 868-881.

(3) Nakamae, K., Miura, K., & Fujioka, H. (1996). VLSI testing with CAD-linked electron beam test system. *Microelectronic Engineering*, *31*(1-4), 319-330.

27. Taken together, the "MAY 13 1993" date stamp on the copyright page of Weste (**Ex. 1014**), the Bibliographic record (**Appendix C**), the MARC record (**Appendix D**), the copyright record (**Appendix E**), and my understanding of the ordinary and customary cataloging and processing practices of libraries inform my opinion that Weste (**Ex. 1014**) would have been cataloged and discoverable by the public at the online catalog of the Library of Congress by April 1992 when the CIP cataloging record was created, and the physical item would have been available for public access by September 16, 1993, three months after the CIP record was verified. A person reasonably familiar with library resources and on-line searching

could have found this book through a reasonably diligent search—by title, authors, series, subject classification numbers, and subject headings—on or before that date. The earliest citing document of Weste appeared in December 1994.

III. CONCLUSION

28. I hereby declare that all statements made herein on my own knowledge are true and that all statements made on information and belief are believed to be true, and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Date: 3-4-2019 Executed: Ingrid Hsieh-Yee

Ingrid Hsieh-Yee, Ph.D.

IV. APPENDICES

APPENDIX A

Ingrid Hsieh-Yee

Professor
Dept. of Library and Information Science
Catholic University of America
Washington, D.C. 20064
E-mail: hsiehyee@cua.edu
Phone: (202) 319-5085
Fax: (202) 319-5574

Education

Ph.D. Library and Information Studies, University of Wisconsin-Madison
Minors: Sociology and Psychology

M.A. Library and Information Studies, University of Wisconsin-Madison.

M.A. Comparative Literature, University of Wisconsin-Madison.

B.A. Foreign Languages and Literature, National Taiwan University.

Work Experience

Professor, School/Dept. of Library and Information Science, Catholic University of America, 2004- (Assistant Professor, 1990-1996; Associate Professor, 1997-2004)

Co-Chair, Dept. of Library and Information Science, Catholic University of America, June 2015-August 2016.

Acting Dean, School of Library and Information Science, Catholic University of America, January 2010-June 2012.

Cataloger, Dept. of Legislative Reference Library, Annapolis, Maryland, 1989-1990.

Lecturer, School of Library and Information Studies, University of Wisconsin-Madison, 1988.

Teaching Assistant, School of Library and Information Studies, University of Wisconsin-Madison, 1986-1988.

Cataloger, Health Sciences Library, University of Wisconsin-Madison, 1984-1986.

Areas of Teaching and Research Interests

Information Organization and Access; Metadata; Cataloging & Classification; Information Architecture; Information Retrieval; Digital Collections; Scholarly Communication; Information

Behavior; Health Informatics; Human Computer Interaction; Usability Studies

Grants & Honors

Cultural Heritage Information Management Project. IMLS grant. Amount: $498,741. Period: Aug. 2012 to July 2015. Co-PI with Dr. Youngok Choi.

D.C. Health Information Technology (HIT4): Building Capacity & Providing Access in Our Nation's Capital. Dept. of Labor H2B Training Grant. Grant amount: $4,175,500. Grant period: Nov. 2011 to Dec. 2015. Partner with the Metropolitan School of Professional Studies of the Catholic University of America, Children's National Medical Center, D.C. Department of Employment Services, Holy Cross Hospital, Howard University, Center for Urban Progress, Providence Hospital, and Sibley Memorial Hospital.

Capital Health Careers Project. Department of Labor Healthcare Sector and Other High Growth and Emerging Industries Grant. Grant amount: $4,953,999. Grant period: March 2010 – February 2013. Awarded to a group of healthcare organizations and educational institutions in Washington, D.C. Providence Health Foundation of Providence Hospital (Lead institution). Part of the grant supported the development of a Master's degree program in Information Technology with a concentration in Health Information Technology offered by the School of Library and Information Science.

The Washington D.C. School Librarians Project. IMLS grant. Grant amount: $412,660. Grant period: Aug. 2007 – June 2011. The School partnered with the District of Columbia Public Schools (DCPS) and the District of Columbia Library Association to educate and mentor school media specialists for the DCPS system. PI, Jan. 2010 to June 2011.

SIG Member of the Year, American Society for Information Science and Technology (2009).

Most Outstanding Paper of *OCLC Systems & Services* (2001).

ALISE Research Grant (2001).

Most Outstanding Paper of *OCLC Systems & Services* (2000).

Research Grant from ERIC (1999-2000).

Best Research Paper Award; Association for Library and Information Science Education (1998).

Research Grants, Catholic University of America. 1991, 1992, 1993, 1996, 1998, 1999, 2004, 2005, 2006, 2007, 2013-14.

Cooperative Faculty Research Grant, Consortium of Universities in the Washington Metropolitan Area (1993-1994).

Cooperative Research Grant, Council on Library Resources (1993-1994).

Journal of the American Society for Information Science Best Paper Award (1993).

ASIS/ISI Information Science Doctoral Dissertation Scholarship (1989).

HEA Title IIB Fellowship (Dept. of Education) (1989)

Chinese-American Librarians Association Scholarship (1987).

Beta Phi Mu (1985).

Vilas Fellowship, University of Wisconsin-Madison. 1984

Publications

Choi, Y., and Hsieh-Yee, I. (2010). Finding Images in an OPAC: Analysis of User Queries, Subject Headings, and Description Notes. *Canadian Journal of Information and Library Science,* 34(3): 271 – 295.

Hsieh-Yee, I. (2008). Educating Cataloging Professionals in a Changing Information Environment. *Journal of Education for Library and Information Science,* 46(2): 93-106.

Vellucci, S. L., Hsieh-Yee, I., and Moen, W.E. (2007). The Metadata Education and Research Information Commons (MERIC): A Collaborative Teaching and Research Initiative. *Education for Information,* 25(3&4): 169-178.

NISO Framework for Guidelines for Building Good Digital Collections. 3rd ed. Baltimore, MD: National Information Standards Organization, 2007. Also available online: http://www.niso.org/framework/framework3.pdf (NISO Working Group members: Priscilla Caplan (chair), Grace Agnew, Murtha Baca, Tony Gill, Carl Fleischhauer, Ingrid Hsieh-Yee, Jill Koelling, and Christie Stephenson.)

Choi, Y., Hsieh-Yee, I., and Kules, B. (2007). Retrieval Effectiveness of TOC and LCSH. *Proceedings of the Joint Conference on Digital Libraries*, pp. 233-234.

Vellucci, S., and Hsieh-Yee, I. (2007). They Didn't Teach Me That in Library School! Building a Digital Teaching Commons to Enhance Metadata Teaching, Learning and Research. *Proceedings of the National Conference of the Association of College and Research Libraries, Baltimore, MD*, pp. 26-31.

Mitchell, Vanessa, and Ingrid Hsieh-Yee. (2007). Converting Ulrich's Subject Headings to FAST Headings: A Feasibility Study. *Cataloging & Classification Quarterly,* 45(1): 59-85.

Hsieh-Yee, I., Tang, R., and Zhang, S. (2007). User Perceptions of a Federated Search System. *IEEE Technical Committee on Digital Libraries Bulletin,* Summer 3(2) (URL = http://www.ieee-tcdl.org).

Tang, R., Hsieh-Yee, I., and Zhang, S. (2007). User Perceptions of MetaLib Combined Search: An Investigation of How Users Make Sense of Federated Searching." *Internet Reference Services Quarterly,* 12(12): 211-236.

Hsieh-Yee, I., Tang, R., and Zhang, S. (2006). User Perceptions of a Federated Search System. *Proceedings of the Joint Conference on Digital Libraries*, *June 11-15, 2006, Chapel Hill*, p. 338.

Hsieh-Yee, I. (2006). *Organizing Audiovisual and Electronic Resources for Access: A Cataloging Guide*. 2nd ed. Westport, Conn.: Libraries Unlimited.

NISO A Framework of Guidance for Building Good Digital Collections. 2nd ed. Bethesda, MD: National Information Standards Organization, 2004. Framework Advisory Group: Grace Agnew, Liz Bishoff, Priscilla Caplan (Chair), Rebecca Gunther and Ingrid Hsieh-Yee.

Hsieh-Yee, I. (2004). Cataloging and Metadata Education in North American LIS Programs. *Library Resources & Technical Services,* 48(1): 59-68.

Hsieh-Yee, I. (2004). Cataloging and Metadata Education. In Gary E. Gorman (Ed.), *International Yearbook of Library and Information Management 2003: Metadata Applications and Management*, (pp.204-234). London: Facet Publishing.

Yee, P. L., Hsieh-Yee, I., Pierce, G.R., Grome, R., and Schantz, L. (2004). Self-Evaluative Intrusive Thoughts Impede Successful Searching on the Internet. *Computers in Human Behavior,* 20(1): 85-101.

Hsieh-Yee, I. (2003). Cataloging and Metadata Education: A Proposal for Preparing Cataloging Professionals of the 21st Century. A report submitted to the ALCTS-Education Task Force in response to Action Item 5.1 of the *Bibliographic control of Web Resources: A Library of Congress Action Plan*. Approved by the Association for Library Collections and Technical Services. Web version available since April 2003 at http://lcweb.loc.gov/catdir/bibcontrol/CatalogingandMetadataEducation.pdf.

Hsieh-Yee, I. (2002). Cataloging and Metadata Education: Asserting a Central Role in Information Organization. *Cataloging & Classification Quarterly* 34(½): 203-222.

Hsieh-Yee, I., and Smith, M. (2001). The CORC Experience: Survey of Founding Libraries, Part I. *OCLC Systems & Services,* 17: 133-140. (Received "The Most Outstanding Paper of OCLC Systems & Services in 2001" award.)

Hsieh-Yee, I., and Smith, M. (2001). The CORC Experience: Survey of Founding Libraries, Part II, Automated Tools and Usage. *OCLC Systems & Services,* 17: 166-177. (Received "The Most Outstanding Paper of OCLC Systems & Services in 2001" award.)

Hsieh-Yee, I. (2001). ERIC User Services: Changes and Evaluation for the Future. *Government Information Quarterly,* 18: 31-42.

Hsieh-Yee, Ingrid. (2001). Research on Web Search Behavior. *Library and Information Science Research,* 23: 167-185.

Logan, E., and Hsieh-Yee, I. (2001). Library and Information Science Education in the Nineties. *Annual Review of Information Science and Technology,* 35: 425-477.

Hsieh-Yee, I. (Ed.) (2001). *Library and Information Science Research*, 23 (2). A special issue in honor of the retirement of Douglas L. Zweizig.

Hsieh-Yee, I. (2000). *ERIC User Services: Evaluation in a Decentralized Environment*. Washington, D.C.: Dept. of Education.

Hsieh-Yee, Ingrid. (2000). *Organizing Audiovisual and Electronic Resources for Access: A Cataloging Guide*. Littleton, CO: Libraries Unlimited.

Hsieh-Yee, I. (2000). Organizing Internet Resources: Teaching Cataloging Standards and Beyond. *OCLC Systems & Services,* 16: 130-143. (Received "The Most Outstanding Paper of OCLC Systems & Services in 2000" award.)

Hsieh-Yee, I. (1998). The Retrieval Power of Selected Search Engines: How Well Do They Address General Reference Questions and Subject Questions? *Reference Librarian,* 60: 27-47.

Hsieh-Yee, I. (1998). Search Tactics of Web Users in Searching for Texts, Graphics, Known Items and Subjects: A Search Simulation Study. *Reference Librarian,* 60: 61-85. (Received the 1997 Best ALISE Research Paper Award.)

Hsieh-Yee, I. (1997). Access to OCLC and Internet Resources: LIS Educators' Views and Teaching Practices. *RQ,* 36: 569-86.

Hsieh-Yee, I. (1997). Teaching Online and CD-ROM Resources: LIS Educators' Views and Practices. *Journal of Education for Library and Information Science,* 38: 14-34.

Hsieh-Yee, I. (1996). The Cataloging Practices of Special Libraries and Their Relationship with OCLC. *Special Libraries,* 87: 10-20.

Hsieh-Yee, I. (1996). Modifying Cataloging Practice and OCLC Infrastructure for Effective Organization of Internet Resources. In *Proceedings of the OCLC Internet Cataloging Colloquium*. [Online]. Available: http://www.oclc.org/oclc/man/colloq/hsieh.htm

Hsieh-Yee, I. (1996). Student Use of Online Catalogs and Other Information Channels. *College & Research Libraries,* 57: 161-175.

Hsieh-Yee, I. (1995). Ten entries in James S. C. Hu (Ed.), *Encyclopedia of Library & Information Science*, 913, 1028-29, 1036, 1037, 1145-46, 1514, 1575, 1763-64, 2216-27, 2378-79. Taipei, Taiwan: Sino-American Publishing. (Topics include "Advanced Technology/Libraries," "Information Ethics," "Instruction on Cataloging and Classification," "Instruction on Reference Services.")

Hsieh-Yee, I. (1993). Effects of Search Experience and Subject Knowledge on Online Search Behavior: Measuring the Search Tactics of Novice and Experienced Searchers. *Journal of the American Society for Information Science,* 44: 161-174. (Received the 1993 Best JASIS Paper Award.)

Presentations

Hsieh-Yee, I. and Fragan-Fly, J. (May 2018) Trends, Design & Strategies for Digital Scholarship Services. Presented at the 2018 Maryland/Delaware Library Association Conference, Cambridge, MD.

Hsieh-Yee, I. (February, 2018) Research Data Management: What It Takes to Succeed. Presented at the 10th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. (February, 2017) *Research Data Management: New Competencies and Opportunities for Information Professionals*. Presented at the 9th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. and Lawton, P. (February, 2017) *Enhancing Catholic Portal Searches with User Terms and LCSH*. Presented at the 9th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. (2016, October) *Visualizing Data for Information*. Presented at the 2016 Virginia Library Association Conference, Hot Springs, VA.

Hsieh-Yee, I. (2016, August) *Religious Materials Toolbox for Archivists: Solutions to Problems Facing the Profession*. Presented at Archives * Records 2016, Atlanta, GA.

Hsieh-Yee, I. and Lawton, P. (2016, March) *Enhancing Retrieval of Catholic Materials with LCSH Knowledge Structure*. Presented at the 2016 Catholic Library Association Conference, San Diego, CA.

Fagan-Fry, J. and Hsieh-Yee, I. (2016, February) *Approaches to Digital Scholarship at Top Universities around the World: Scholarly Publishing in the Digital Age.* Presented at the 8th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee and Fagan-Fry, J. (2016, January) *Innovative Services for Digital Scholarship at Top 100 Research Libraries of the World*. Poster presented at the 2016 Annual Conference of the Association for Library and Information Science Education, Boston, Mass.

Hsieh-Yee, I. and Lawton, P. (2015, June). *Crowdsourcing terms for CRRA portal themes.* Poster presented at the third CRRA symposium and annual meeting, Bringing the created toward the Creator: Liturgical art and design since Vatican II. Catholic Theological Union, Chicago, Illinois.

Hsieh-Yee, I. and Lawton, P. (2015, February). *Crowdsourcing terms for thematic exploration in the Catholic Portal.* Poster presented at the 7th Annual Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I., James, R., and Fagan-Fry, J. (2015, February). *Support for digital scholarship at top university libraries of the world*. Poster presented at the 7th Annual Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I., Zhang, S., Lin, K., and Cherry, S. (2015, February). *Thus said the end users: Summon experience and support for research workflows*. Poster presented at the 7th Annual Bridging the Spectrum Symposium, Washington, D.C.

Yontz, E., Hsieh-Yee, I., & Houston, S. (2015, February). *Healthy Heroes Summer Reading Club: Developing healthy youth at public libraries.* 11th Annual Jean Mills Health Symposium, Greenville, North Carolina.

Yontz, E., Hsieh-Yee, I., and Houston, S. (2015, January). *Healthy youth and libraries: A pilot study.* Association for Library & Information Science Education (ALISE) Annual Conference, Chicago, Illinois.

Hsieh-Yee, I. (2014, May). *Linking CRRA resources to portal themes via authority files*. Presented at the Catholic Research Resources Alliance 2014 Membership Meeting, Marquette, WI.

Hsieh-Yee, I. (2014, April). *Enhancing subject access to CRRA resources*. Presented at the 2014 Catholic Library Association Conference, Pittsburgh, PA.

Hsieh-Yee, I. (2014, January). *Health Information Technology Program: Educational entrepreneurship in action*. Presented at the 2014 annual Conference of the Association for Library and Information Science Education, Philadelphia, PA

Hsieh-Yee, I., Zhang, S., Lin, K., and Cherry, S. (2014, January). *Discovering information through Summon: An analysis of user search strategies and search success*. Paper presented at the 6th Bridging the Spectrum Symposium, Washington, D.C.

Hsieh-Yee, I. (2012, December). *National Digital Stewardship Alliance and SLIS at CUA: An Educational Partnership*. Paper presented at Best Practices Exchange: Acquiring, Preserving, and Providing Access to Government Information in the Digital Era, Annapolis, MD

Choi, Y. and Hsieh-Yee, I. (2010, January). *Finding Images in an OPAC: Analysis of User Queries, Subject Headings, and Description Notes*. Paper presented at 2nd Annual Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. and Coogan, J. (2010, January). *Google Scholar vs. Academic Search Premier: What Libraries and Searchers Need to Know*. Paper presented at 2nd Annual Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (2009, November). *Information Science Education: An LIS School's Perspective*. Paper presented at Annual Meeting of the American Society for Information Science and Technology, Vancouver, British Columbia, Canada.

Hsieh-Yee, I., Menard, E., Ya-Ning Chen, A., Shu-Jiun Chen, S., Kalfatovic, M. R., Wisser. K. M. (2009, November). *Information Organization in Libraries, Archives and Museums: Converging Practices and Collaboration Opportunities*. Presented at Annual Meeting of the American Society for Information Science and Technology, Vancouver, British Columbia, Canada. (Organizer and moderator of this panel.)

Hsieh-Yee, I. and Coogan, J. (2009, July). *Catching up to Google Scholar: The Retrieval Power of Academic Search Premier and Google Scholar*. Poster presented at American Library Association Conference, Chicago, Illinois.

Hsieh-Yee, I., with the CUA Scholarly Communications Project Team. (2009, January). *Digital Scholarship@CUA: Developing an Institutional Repository for CUA*. Poster presented at 1st Annual Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Wise, M., Cylke, K., and Hsieh-Yee, I. **(**2009, January). *Digital Talking Books: Meeting the Needs of the Blind and the Handicapped*. Paper presented at the Bridging the Spectrum Symposium, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (2009, January). *User Expectations of MERIC*. Presented at the Information Organization Competencies for the 21st Century Discussion Session of the 2009 Conference of the Association for Library and Information Science Education, Denver, Colorado.

Choi, Y., and Hsieh-Yee, I. (2008,November). *Subject Access for Images in an OPAC.* Annual Meeting of the American Society for Information Science and Technology, Columbus, Ohio. (Also co-organized a panel on Retrieving and Using Visual Resources: Challenges and Opportunities for Research and Education.)

Hsieh-Yee, I. (2008, June). *Educating Cataloging Professionals in a Changing Information Environment*. National Taiwan University, Taipei, Taiwan.

Vellucci, S. L., Moen, W.E., Hsieh-Yee, I., Marson, B., and Wisser, K. (2008, January) *Building a Metadata Education and Research Community through MERIC (Metadata Education and Research Information Commons): Demo and Stakeholder Input*. A panel presented at the 2008 Conference of the Association for Library and Information Science Education, Philadelphia, Pennsylvania.

Hsieh-Yee, I., Choi,Y. and Kules, B. (2007, October). *Searching for Books and Images in OPAC: Effects of LCSH, TOC and Subject Domains.* A poster presented at the American Society for Information Science and Technology Annual Meeting, Milwaukee, Wisconsin.

Hsieh-Yee, I. and Coogan, J. (2007, August) *Google Scholar vs. Academic Search Premier: A Comparative Analysis.* Presented to the Faculty and Staff of the University of the District of Columbia.

Hsieh-Yee, I. and Coogan, J. (2007, June). *Google Scholar vs. Academic Search Premier: A Comparative Analysis*. Presented to the Washington Research Library Consortium Community, Catholic University of America, Washington, D.C.

Hsieh-Yee, I., Choi, Y., and Kules, B.. (2007, June). *What Users Need for Subject Access: Table of Contents or Subject Headings?* A poster presented at the 2007 American Library Association Annual Conference, Washington, D.C., June 2007.

Choi, Y., Hsieh-Yee, I., and Kules, B. (2007, June). *Retrieval Effectiveness of TOC and LCSH.* A paper presented at the Joint Conference on Digital Libraries 2007, Vancouver, Canada.

Vellucci, S. L., Hsieh-Yee, I., and Moen, W.E. (2007, May). *If We Build It, Will They Come? Building a Community of Practice for Metadata Stakeholders.* A poster presented at the Rutgers University Research Day, Bridgeton, New Jersey.

Hsieh-Yee, I. (2007, May). *Federated Searching: User Experience & Perceptions.* International Conference on Information Organization & Retrieval, National Taiwan University, Taipei, Taiwan.

Hsieh-Yee, I. (2007, May). *Search Performance of Google Scholar and Academic Search Premier.* International Conference on Information Organization & Retrieval, National Taiwan University, Taipei, Taiwan.

Hsieh-Yee, I. (2007, May) *MERIC: Building a Digital Commons for Metadata Education & Research.* International Conference on Information Organization & Retrieval, National Taiwan University, Taipei, Taiwan.

Hsieh-Yee, I., and Coogan, J. (2007, March/April). *A Comparative Analysis of Google Scholar and Academic Search Premier.* Poster presented at the Association of College & Research Libraries 13th National Conference, Baltimore, Maryland.

Vellucci, S. L. and Hsieh-Yee, I. (2007, March/April) *They Didn't Teach Me That in Library School! Building a Digital Teaching Commons to Enhance Metadata Teaching, Learning and Research*. On-site presentation and Webcast by Elluminate. A contributed paper presented at the Association of College & Research Libraries 13th National Conference, Baltimore, Maryland. The acceptance rate for contributed paper was 20%. This paper was one of 10 conference papers chosen for live webcast during the conference.

Moen, W., Hsieh-Yee, I. and Vellucci, S.L. (2007, January) *A DSpace Foundation for a Teaching & Research Commons: The Metadata Education and Research Information Commons.* A poster session presented at the Open Repositories Conference 2007, San Antonio, Texas.

Tang, R., Hsieh-Yee, I., and Zhang, S. (2006, November) *User Perception of MetaLib Combined Search*. Paper presented at the Annual Meeting of the American Society for Information Science and Technology, Austin, Texas, Nov. 2006.

Hsieh-Yee, I. (2006, November). *Federated Searching: User Perceptions, System Design, and Library Instructions.* Paper presented at the Annual Meeting of the American Society for Information Science and Technology, Austin, Texas. (Panel organizer, moderator, presenter).

Hsieh-Yee, I. (2006, November). *Building a Digital Teaching Commons to Enhance Teaching and Learning: The MERIC Experience and Challenges.* Paper presented at the Annual Meeting of the American Society for Information Science and Technology, Austin, Texas. (Panel organizer, moderator, presenter)

Hsieh-Yee, I. (2006, September). *Search Performance of Google Scholar and Academic Search Premier.* Paper presented at the ERIC Publishers Meeting, Washington, D.C.

Hsieh-Yee, I., Zhang, S., and Rong Tang, R. (2006, June). *User Perceptions of a Federated Search System*. Poster presented at Joint Conference on Digital Libraries, Chapel Hill, North Carolina.

Hsieh-Yee, I. and Zhang, S. (2006, June). *Preparing Users for Federated Search: Implications of a MetaLib User Perceptions Study.* Paper presented at the 2006 Ex Libris User Groups of North America Conference, Knoxville, Tennessee.

Hsieh-Yee, I. (2006, January). *MERIC Organizations and Navigation*. Paper presented at the 2006 ALISE Annual Conference, San Antonio, Texas.

Hsieh-Yee, I. (2006, January). *Metadata and Cataloging Education: Recommended Competencies.* Paper presented at the 2006 ALISE Annual Conference, San Antonio, Texas.

Hsieh-Yee, I. (2005, November). *Digital Library Evaluation: Progress & Next Steps.* Presentation at the Annual Meeting of the American Society for Information Science & Technology, Charlotte, North Carolina.

Hsieh-Yee, I. (2005, August). *Providing Access to Digital Content: Issues for DL Managers.* Presentation at MDK12 Digital Library Steering Committee Meeting, Columbia, Maryland.

Hsieh-Yee, I. (2005, April). *Enhancing Teaching and Learning: The Role of School Library Media Specialists.* Presentation at Meeting of the Baltimore County Public School System School Media Specialists, Baltimore, Maryland.

Hsieh-Yee, I. (2005, January). *Subject Access and Users: Insights & Inspirations from Marcia J. Bates*. Paper presented at the Historical Perspectives SIG, 2005 Conference of the Association for Library and Information Science Education, Boston, Massachusetts.

Hsieh-Yee, I. (2005, January). *Electronic Resource Management: Practice, Employer Expectations, & CE Interests.* Paper presented at Technical Services Education SIG, 2005 Conference of the Association for Library and Information Science Education, Boston, Massachusetts.

Hsieh-Yee, I. (2004, October). *Library Professionals for the Digital Age: Competencies & Preparation.* Paper presented at Bibliographic Access Management Team meeting, Library of Congress, Washington, D.C.

Hsieh-Yee, I. (2004, January). *Cataloging and metadata expertise for the digital era.* Presented at Preparing 21st Century Cataloging and Metadata Professionals: A Workshop for Educators and Trainers, San Diego and sponsored by ALCTS, ALISE, LC, and OCLC.

Hsieh-Yee, I. (2004, January). *Educating catalogers for the digital era.* Paper presented at the Technical Services SIG, 2004 Conference of the Association for Library and Information Science Education, San Diego.

Hsieh-Yee, I. (2003, July). *Cataloging Education for the 21st Century.* A presentation at the Library of Congress, Washington, D.C.

Hsieh-Yee, I. (2002, January) *Metadata Education and Research Priorities: A Delphi Study of*

Metadata Experts. Presentation at the 2002 Conference of the Association for Library and Information Science Education, New Orleans.

Hsieh-Yee, I. (2001, November). *A Delphi Study of Metadata: Preliminary Findings.* Poster session at the 2001 Annual Meeting of the American Society for Information Science & Technology, Washington, D.C.

Hsieh-Yee, I. (2001, June). *Resources on Asian American Children: Analysis of Retrieval by Search Engines and WorldCat.* Presentation at the National Conference on Asian Pacific American Librarians, San Francisco.

Hsieh-Yee, I. (2001, January). *Delphi Study on Metadata: Project Design.* Presentation at Research Awards Session, Association for Library & Information Science Education, Washington, D.C.

Hsieh-Yee, I. (2000, May). *Web Search Behavior Research: Progress and Implications.* Presentation at the Symposium on Evaluating Library and Information Science Research, University of Wisconsin-Madison, Madison, Wisconsin.

Hsieh-Yee, I. (2000, March). *ERIC User Services: Evaluation in a Decentralized Environment.* Presentation at the National ERIC Joint Directors/Technical Meeting, Arlington, Virginia.

Hsieh-Yee, I. (2000, January). *Enhancing Learning with Web Technology.* Presentation at Faculty Conversations, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (2000, January). *From Surrogates to Objects: CUA's Approaches to Organizing Electronic Resources.* Paper presentation at the Annual Conference of the Association for Library and Information Science Education, San Antonio, Texas.

Yee, P., and Hsieh-Yee, I. (1997, November). *Individual Differences in Search Behavior on the WWW.* A poster session presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia, Pennsylvania.

Hsieh-Yee, I. (1997, April). *Research + Marketing + Preparation = Job!* Presented at the "Workshop on Resume and Interview Techniques," Special Libraries Association, Student Chapter, Catholic University of America, Washington, D.C.

Hsieh-Yee, I. (1997, February). *Creating CyberCatalogers: Education and Training.* Presentation at ALA's Midwinter Meeting, Washington, D.C.

Hsieh-Yee, I. (1997, February). *Search Tactics of Web Users in Searching for Texts, Graphics, Known Items and Subjects: A Search Simulation Study.* Presented at the Conference of the Association for Library and Information Science Education, Washington, D.C.

Hsieh-Yee, Ingrid. "Beginning Your Special Library/Information Center Career." Presented at SLA's "Career Day," Jan. 11, 1997, Catholic University of America.

Hsieh-Yee, I. (1996, September). *The Roles of Library and Information Scientists in Managing Electronic Information.* Presentation at Hamilton College, Clinton, New York.

Hsieh-Yee, I. (1996, May). *The Future of Cataloging as a Profession.* Presented at "The Cataloging Forum, Library of Congress, Washington, D.C.

Hsieh-Yee, I. (1994, October). *The Impact of the Internet on OPACs.* Presented at the Third Workshop on User Interfaces for OPACs, Library of Congress, Washington, D.C.

Reports

Hsieh-Yee, I., with Knowledge Management Competencies and Performance Action Group of the Federal Knowledge Management Initiative. "From Knowledge Management Competencies to Improved Organizational Performance." April 9, 2009.

Hsieh-Yee, I., with Knowledge Practices Action Group of the Federal Knowledge Management Initiative. "KM Practice in Government Agencies: Findings and Recommendations." April 9, 2009.

Hsieh-Yee, I. "Delphi Study on Metadata." 2001. Three quarterly reports submitted to the Association for Library and Information Science Education.

Hsieh-Yee, I. "College Students' Information Channels: Patterns of Use and Possible Factors in Channel Selection." 1995. Submitted to the Catholic University of America.

Hsieh-Yee, I. "The Information-Seeking Patterns of Scholars and Their Use of an Online Information System." 1994. Submitted to the Council on Library Resources.

Book Reviews

Review of *The Measurement and Evaluation of Library Services*, by Sharon L. Baker and F. Wilfrid Lancaster. *Information Processing and Management* 30 (1994): 450-52.

Review of *Subject Access to Films and Videos, by Sheila S. Intner and William E. Studwell; and Cataloging Unpublished Nonprint Materials*, by Verna Urbanski with Bao Chu Chang and Bernard L. Karon. *Information Processing and Management* 30 (1994): 449-50.

Review of *Automated Information Retrieval in Libraries: A Management Handbook*, by Vicki Anders. *Journal of Library and Information Science* 19 (1993): 98-100.

Review of *Full Text Databases*, by Carol Tenopir and Jung Soon Ro. *Information Processing and Management* 28 (1992): 667-68.

Review of *Descriptive Cataloging for the AACR2R And USMARC: A How-to-Do It Workbook*, by Larry Millsap and Terry Ellen Ferl. *Information Processing and Management* 28 (1992): 809-11.

Review of *MARC Manual: Understanding and Using MARC Records*, by Deborah J. Byrne. *Information Processing and Management* 28 (1992): 537-38.

Service

Professional Associations and Societies

- Library of Congress. RDA Training Program for the Profession. Co-authored with Tim Carlton. 2013-2014.
- 2014 Digital Preservation Outreach & Education Survey. Contributed to the design of the survey, 2014.
- National Digital Stewardship Alliance. Outreach Committee. 2011-2014.
- National Digital Stewardship Residency Program. Advisory Group, 2012-2013.
- FEDLINK Health Information Technology Advisory Council, 2011-2015.
- 2012 Joint Conference on Digital Libraries. Program Planning Committee, Pre-Conference Proposals Review Committee, 2012
- Catholic Research Resources Alliance. Five-Year Strategic Plan Task Force, 2011-2012
- Institute of Museum and Library Services. Grant reviewer. 2004, 2005, 2010.

- Association for Library and Information Science Education.
  * ALISE Bodan Wynar Research Paper Award Committee, 2015, 2016, 2017
  * ALISE Eugene Garfield Dissertation Award Competition, Jury, 2013, 2014
  * ALISE Research Grant Competition Committee. Chair, 2012
  * Pratt-Severn Faculty Innovation Award. Chair, 2009, 2010
  * ALISE Doctoral Poster Jury, 2012
  * "Information Organization Competencies for the 21st Century" Discussion session leader. 2009 Conference of the Association for Library and Information Science Education.
  * Assisted Technical Services SIG Convener in organizing a program, ""Building a Metadata Education and Research Community through MERIC (Metadata Education and Research Information Commons): Demo and Stakeholder Input" for the 2008 ALISE conference.
  * Association for Library Collections and Technical Services/Association for Library and Information Science Education (ALCTS/ALISE) Metadata Education and Research Information Center (MERIC) Advisory Board, Co-Chair (with Sherry Vellucci), 2005-2007. Chair, 2008-2009 (leading the effort to build MERIC, a repository and collaborative space for metadata educators, practitioners, and researchers)

* Technical Services SIG, Convener, 2004-2005. Organized a program on “Electronic Resources Management: Current Practices, Employer Expectations, and Teaching Strategies” for the 2005 conference in Boston, Massachusetts.
* Technical Services SIG, Convener, 2003-2004. Organized a program on “Organizing Information with Metadata: Desired Competencies and Teaching Innovations” for the 2004 conference.
* Technical Services SIG, Convener, 1999-2000. Organized a program on "Teaching the Organization of Electronic Resources" for the 2000 conference.
* Curriculum SIG, Co-convener (with Sibyl Moses), 1996-97. Organized a program on “Government Information Policy” for the 1997 conference.

- American Society for Information Science & Technology.
  * Reviewer, Conference program panel submissions and poster submissions, 2005, 2006, 2007, 2009, 2011, 2012, 2013, 2014, 2015, 2016, 2017
  * Nomination Committee, 2009-2011
  * Information Science Education Special Interest Group. American Society for Information Science and Technology. Chair-Elect, 2007-2008. Chair 2008-2009.
  * Committee on Information Science Education. 1999-2006.
  * Committee on Information Science Education. Organizing Committee for an orientation program for students at ASIS annual meetings, 1999-2001
  * Committee on Information Science Education. Sub-committee on Student Welfare (focusing on issues related to master's education), 1998-2001
  * SIG ED. Organizing Committee for the "Seminar on Research and Career Development" for junior researchers. 1995-96 (chair), 1997-2001
  * ISI Doctoral Dissertation Proposal Scholarship Jury, 1997; 2001, 2002
  * Pratt-Severn Best Student Research Paper Award Jury. Chair. 1997
  * 1998 Midyear Meeting (referee of contributed papers), 1997
  * Organizer and moderator of the ASIS Doctoral Forum and the Doctoral Research Seminar 1994-1995
  * SIG Human Computer Interaction. Chair-Elect, Chair, 1993-1995
  * Doctoral Forum Award Jury, 1995
  * Best Student Paper Award Jury, 1995

- American Library Association.
  * Committee on Accreditation, External Review Panelist, 2009- (site visiting team 2013-2014; site visiting team 2016-2017)
  * Association for Library Collections and Technical Services Task Force on Competencies and Education for a Career in Cataloging, member, 2008-2009
  * Facilitator for “What They Don't Teach in Library School: Competencies, Education and Employer Expectations for a Career in Cataloging,” an Association for Library Collections and Technical Services Preconference, June 22, 2007 in Washington, D.C. Also a local liaison for bringing this program to the Catholic University of America.
  * Facilitator for a discussion on "Effect of Electronic Resources on Technical Services" at ALA's Midwinter Meeting held in Feb. 1997 in Washington, D.C.

* International Relations Committee, Subcommittee Task Force for IFLA and China, 1994-1997

- Virginia Association of School Librarians. Scholarships and Awards Committee. 2010-2012

- Federal Knowledge Management Initiative, Knowledge Management Practices Action Group. Member. 2009 (leading the effort to build a knowledge management repository)
- Federal Knowledge Management Initiative, Knowledge Management Competencies & Learning Action Group. Member. 2009 (developing an action plan for helping government knowledge workers and government agencies to develop knowledge management competencies)
- National Center for Education Statistics. Technical Review Panel. 2008.
- External evaluator for a case of promotion to full professorship. University of Tennessee. 2008.
- National Information Standards Organization (NISO). Advisory Board, Revision of "IMLS Framework of Guidance for Building Good Digital Collections," 2004, 2007.
- Library of Congress, Bibliographic Control of Web Resources: A Library of Congress Action Plan. Principal Investigator of Action Item 5.1, focusing on cataloging and metadata education for students and new librarians, 2002-2003. (worked with the Association for Library Collections and Technical Services, Education Task Force)

- Chinese American Librarians Association
  * Chinese American Librarians Association Outstanding Library Leadership Award in Memory of Dr. Margaret Chang Fung, Award Committee, 2016-2017
  * Achievement Award Jury, 2000-2001
  * CALA Goal 2000 Task Force, 1997
  * Scholarship Committee, 1995, 1996-1997 (chair)
  * Board of Directors, 1994-1997
  * Publication Committee, 1993-1995
  * International Relations Committee, 1993-1996

- Sailor[SM] Assessment Advisory Group (An impact study of Sailor, Maryland's Public Information Network), 1995

- Editorial boards
  Journal of Library and Information Science. Editorial Board, 2012-
  Chinese American Librarians Association, *Occasional Papers Series*. Editorial Board, 2009-2016.
  Library Quarterly. Editorial Board, 2003-2008
  Bulletin of the Medical Library Association, 1994-97
  Newsletter editor for the Chinese American Librarians Association, 1989-92

- Referee for the following journals
  Information Processing and Management

Journal of Digital Information
Journal of Education for Library and Information Science
Journal of Information Science
Journal of Library & Information Science
Journal of Library Metadata
Journal of the American Society for Information Science & Technology
Library and Information Science Research
Library Quarterly

- Expert reviewer, "Digital Library" course, Evaluation module, University of North Carolina, Chapel Hill, 2007-2008.
- Expert reviewer, "Information Organization" course, University of Michigan, Ann Arbor. 2007.

Catholic University of America

- School of Arts & Sciences, Academic Senate representative, 2017-2020
- School of Arts & Sciences, Committee on Appointments and Promotions, 2015-2019
- School of Arts & Sciences, Academic Council, 2015-2016.
- School of Arts & Sciences, Ordinary Professor Group, 2013-
- Doctoral Dissertation Defense Committee, Chair, Dept. of Psychology, 2016, 2017, 2018
- Doctoral Dissertation Defense Committee, Chair, Dept. of Education, 2014, 2015, 2017
- President's Administrative Council, 2010-2012
- Deans' Council, 2010-2012
- Academic Leadership Group, 2010-2012
- Academic Senate, 2003-2012
- Academic Senate, Committee on Committees and Rules, 2009-2012
- Academic Senate, Committee on Appointments and Promotions, 2005-2008
- Graduate Board, 2010-2012
- CUA Scholarly Communication Project Team, Member (2007), Chair, 2008-2009
- Academic Senate Library Committee, Interim Chair (2007), Member, 2008-2012
- Doctoral Dissertation Defense Committee, Chair, School of Nursing, 2006, 2008
- Dean Search Committee, 1992-1994, 1998-1999, 2002-2003, 2006-2007
- Fulbright Review Panel, 2006
- Academic Senate Committee on Computing, 1995-2003
- CUA Service Learning Advisory Board, 2001-2002
- CUA Faculty Conversations on Enhancing Teaching and Learning through Technology, Planning Group, 1999-2001
- CUA Initiative on Technology and Teaching, 1998-2001

Dept. of Library and Information Science

- Symposium and Colloquium Committee, fall 2016-May 2018, Chair, May 2018-

- Admissions Committee, 2007-2009, Chair 2010-2012, Member 2013-2015, Member 2018-
- Accreditation presentation, Chair, June 2015-August 2016
- Interim Co-Chair, June 2015-August 2016.
- Appointments and Promotions Committee, 1991-
- Blended/OWL Learning Committee, spring 2016-2018
- Scholarship and Awards Committee, fall 2016-
- Technology Committee, fall 2016-2017
- Comprehensive examination editor, 2016-2017, reader (every year since 1990)
- LIS Advisory Board, 2015-2016 (chair); fall 2016- May 2018 (member)
- Committee on Planning and Assessment, 2015-2016 (chair)
- Senior Faculty Committee, 2014-2016.
- Accreditation Steering Committee, 2014-2016 (Chair, 2015-2016)
- Accreditation Students Standard Committee, co-chair, 2014-2016
- Accreditation Mission, Goals, and Objectives Standard Committee, co-chair, 2014-2016
- Accreditation Curriculum Standard, member 2014-2-16
- Accreditation Administration and Finance Standard, member 2014-2016
- Cultural Heritage Information Management Project (IMLS-funded), Co-PI, 2012-2015
- Cultural Heritage Information Management Forum (scheduled for June 2015), Co-Organizer, 2013-2015
- Health Information Technology Interim Review Committee, 2015 (chair)
- Health Sciences Librarianship Advisory Group, 2015- (chair)
- Comprehensive examination editors, 2013-2014, 2016-2017
- National Digital Stewardship Alliance liaison, 2011-2014
- Advisory Board, Chair 2010-2012
- Academic Honesty Committee, Chair, 2008-2012
- Blended Learning Committee, 2010-2012
- Colloquium Committee, 2010-2012
- Comprehensive Examination Administration, 2010-2012
- Cultural Heritage Information Management Advisory Committee, 2010-2012 (chair), 2013-
- Curriculum Committee, 1991-2003, 2007-2009, Chair 2010-2012, member 2013-
- Curriculum Subcommittee on Comprehensive Examination, Chair 2009-2012
- Health Information Technology Advisory Board, Chair 2010-2012. Member 2013-
- Health Sciences Advisory Committee, 2009, Chair 2010-2012. Member 2013-
- HIT Expert Forum, Chair 2012. Member 2013-
- Health Information Technology Student Group Advisor, 2011-2012
- State Council for Higher Education of Virginia, SLIS Representative, 2010-2012
- Symposium Planning Committee, 2010-2012
- Website Management Team, Chair, 2010-2012
- Urban School Librarianship Project (IMLS-Funded), PI, 2007-2011 (chair, 2010-11)
- Failing Grades Committee, 1995-1997 (chair), 2000-2001 (chair), 2004-2005 (chair), 2007 (chair)-2011

- Faculty Search Committee, 1994-1998, 2002-2004, 2006 (chair), Fall 2007-2009, Chair fall 2009-2012
- Recruitment Committee, Chair 2010-2012
- Strategic Planning Committee, Chair 2010-2012
- Technology Committee, 2010-2012
- Accreditation Advisory Committee, 2007-2009
- Accreditation Coordinating Committee, 2007-2009
- Accreditation Steering Committee, 2007-2009
- SLIS Advisory Group, 2007-2009
- Accreditation Curriculum Standard Committee, Co-chair, 2007-2009
- Accreditation Faculty Standard Committee, Co-chair, 2007-2009
- LSC 551 Information Organization Review Team, Co-chair, 2008-2009, 2015-2016.
- Curriculum Subcommittee on Portfolios, 2009
- LSC 555 Information Systems in Libraries and Information Centers Review Team, contributor, 2008-2009
- Redesign of LSC 730 Use and Users of Libraries and Information. 2009-
- Development of a metadata institute that was taught as LSC 715 Organization of Internet Resources in 2008. The institute is being revised and will be offered in 2010 under a new course title.
- Development of lesson plans, assignments, and evaluation rubrics for LSC 606, Cataloging and Classification, for the School's NCATE accreditation. 2008
- Howard and Mathilde Rovelstad Scholarship Committee, Chair, 2004-2007
- Assistant Dean Search Committee, Chair, Fall 2007
- Liaison to the Association for Library Collections and Technical Services to bring its preconference program, Cataloging Education and Employer Expectations, to CUA during the 2007 American Library Association Annual Meeting in Washington, D.C.
- Organizer of the colloquium presentation and reception for Tamar Sadeh of Ex Libris on PRIMO June 2007
- Practicum review and design (work with potential supervisors, such as the American Indian Museum internship description revision) 2006-
- Comprehensive examinations (edits, proctoring, and grading), 1990-
- SLIS Web site redesign: Comments and suggestions. Fall 2007
- Conducted surveys of current students and alumni in preparation for the 2005 re-accreditation, 2004-2005
- Student advisement, 1990-
- Technology Committee, 1992-1999 (chair, 1996-1998), 2002-2003 (member)
- Colloquia Committee 1997-1999, 2002-2003.
- Advisor of the CUA Student Chapter of the American Society for Information Science and Technology, 2002-2003
- Visiting Professor Search Committee, 1999, 2000, 2001
- Leader, Participation in the CORC experiment, 1999-2000
- Advisor of the Special Libraries Association Student Chapter, 1993-1999; the group was recognized for outstanding leadership by SLA in 1999.

- COA planning Committee, Task Force on Electronic Presentation of SLIS Reports (team leader) 1997-1998
- COA Planning Committee, Subcommittee on Technology 1996-1998
- NLM practicum coordinator, 1997-1998
- Computer Literacy Workshops: Assisted with the development and evaluation of the workshops, 1996-1998
- Leader, Participation in the InterCat project, 1995-1997

APPENDIX B

MARC 21 Reference Materials

Part VII:

A Summary of Commonly Used MARC 21 Fields

This is a summary of the MARC 21 tags used most frequently by libraries in entering their own bibliographic records. For full listings of all MARC 21 tags, indicators, and subfield codes, see *MARC 21 Format for Bibliographic Data*.

In the explanations on these pages:

Tags -- The tags (3-digit numbers) are followed by the names of the fields they represent. In this summary, and in the *MARC 21 Format for Bibliographic Data*, if a tag can appear more than once in one bibliographic record, it is labeled repeatable (R). If it can only be used once, it is labeled non-repeatable (NR). For example, a catalog record can have several subjects, so the tags for subject added entries (6XX) are labeled repeatable (R).

Indicators -- The use of indicators is explained in fields where they are used. Indicators are one-digit numbers. Beginning with the 010 field, in every field -- following the tag -- are two character positions, one for Indicator 1 and one for Indicator 2. The indicators are not actually defined in all fields, however. And it is possible that a 2nd indicator will be used, while the 1st indicator remains undefined (or vice versa). When an indicator is undefined, the character position will be represented by the character # (for blank space).

Subfield codes -- All the data in each field (beginning with the 010 field) is divided into subfields, each of which is preceded by a delimiter-subfield code combination. The most common subfield codes used with each tag are shown. Each subfield code is preceded by the character $, signifying a delimiter. The name of the subfield follows the code.

In general, every field MUST have a subfield 'a' (**$a**). One exception that is often seen is in Field 020 (ISBN), when the ISBN information (subfield **$a**) is unavailable but the price (subfield **$c**) is known. Some subfields are repeatable. In this summary, repeatability is noted for only the more common repeatable subfields.

Examples: Examples follow the explanation for each field. For clarity, one space has been placed between the tag and the first indicator, one space has been placed between the

second indicator and the first delimiter- subfield code, and one space has been inserted between the delimiter-subfield code and the subfield data.

010 Library of Congress Control Number -- (LCCN)
(NR, or Not Repeatable
Indicators undefined.
Subfield used most often:
$a -- Library of Congress control number

```
Example:    010 ##  $a ###86000988#
```

020 International Standard Book Number -- (ISBN)
(R, or Repeatable)
Indicators undefined.
Subfields used most often:
$a -- International Standard Book Number
$c -- Terms of availability (often a price)
$z -- Cancelled/invalid ISBN (R)

```
Example:    020 ##  $a 0877547637
```

040 Cataloging source -- (NR)
Indicators undefined.
Subfields used most often:
$a -- Original cataloging agency
$c -- Transcribing agency
$d -- Modifying agency (R)

```
Example:    040 ## $a DLC
                   $c DLC
                   $d gwhs
```

100 Main entry -- Personal name -- (primary author)
(NR; there can be only one main entry)
Indicator 1: Type of personal name entry element
0 -- Forename
1 -- Surname (this is the most common form)
3 -- Family name
Indicator 2 undefined.

Indicator 2 became obsolete in 1990. Older records may display 0 or 1

Subfields used most often:

$a -- Personal name
$b -- Numeration
$c -- Titles and other words associated with a name (R)
$q -- Fuller form of name
$d -- Dates associated with a name (generally, year of birth)

```
Example:    100 1# $a Gregory, Ruth W.
                   $q (Ruth Wilhelme),
                   $d 1910-
```

130 Main entry -- Uniform title -- (NR)

Indicator 1: Nonfiling characters

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

Indicator 2 undefined.

Indicator 2 became obsolete in 1990. (See 100 above.)

Subfields used most often:

$a -- Uniform title
$p -- Name of part/section of a work (R)
$l -- Language of a work
$s -- Version
$f -- Date of a work

```
Example:    130 0# $a Bible.
                   $p O.T.
                   $p Psalms.
```

240 Uniform title (NR)

Indicator 1: Uniform title printed or displayed

0 -- Not printed or displayed
1 -- Printed or displayed (most common)

Indicator 2: Nonfiling characters

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

Subfields used most often:

$a -- Uniform title
$l -- Language of a work
$f -- Date of a work

```
Example:    240 10 $a Ile mystérieuse.
                   $l English.
                   $f 1978
```

245 Title Statement (NR)

Indicator 1: Title added entry

(Should the title be indexed as a title added entry?)

0 -- No title added entry

(indicates a title main entry; i.e. no author is given)

1 -- Title added entry

(the proper indicator when an author given in 1XX; the most common situation)

Indicator 2: Nonfiling characters

0-9 -- Number of nonfiling characters present, including spaces; usually set at zero, except when the title begins with an article; e.g., for *The robe*, the second indicator would be set to 4. The letters *T*, *h*, *e*, and the space following them are then ignored in alphabetizing titles. The record will be automatically filed under "*r*" -- for *Robe*.

Subfields used most often:

$a -- Title proper

$h -- Medium (often used for non-book media)

$p -- Name of part/section of a work (R)

$b -- Reminder of title (subtitles, etc.)

$c -- Remainder of title page transcription/Statement of responsibility

```
Example:    245 14 $a The DNA story :
                   $b a documentary history of gene
                      cloning /
                   $c James D. Watson, John Tooze.
```

246 Varying form of title (R)

Indicator 1: Note/title added entry controller

1 -- Note, title added entry

3 -- No note, title added entry

Indicator 2: Type of title

-- No information provided

0 -- Portion of title

1 -- Parallel title

4 -- Cover title

8 -- Spine title

Subfield used most often:

$a -- Title proper

```
Example:    246 3# $a Four corners power review
```

250 Edition statement (NR)

Indicators undefined.

Subfield used most often:

$a -- Edition statement

```
Example:    250 ## $a 6th ed.
```

260 Publication, distribution, etc. (Imprint) (R)

Indicator 1: Sequence of publishing statements

-- No information provided

Indicator 2: Undefined

Subfields used most often:

$a -- Place of publication, distribution, etc. (R)

$b -- Name of publisher, distributor, etc. (R)

$c -- Date of publication, distribution, etc. (R)

```
Example:    260 ## $a New York :
                   $b Chelsea House,
                   $c 1986.
```

300 Physical description (R)

Indicators undefined.

Subfields used most often:

$a -- Extent (number of pages) (R)

$b -- Other physical details (usually illustration information)

$c -- Dimensions (cm.) (R)

$e -- Accompanying material (for example, "teacher's guide" or "manual")

```
Example:    300 ## $a 139 p. :
                   $b ill. ;
                   $c 24 cm.
```

440 Series statement / Added entry--Title

This field was made obsolete in 2008 to simplify the series statement. See 490 and 830.

490 Series statement (No added entry is traced from field) (R)

Indicator 1: Specifies whether series is traced (whether an 8XX tag is also present)

0 -- Series not traced

1 -- Series traced (8XX is in record)

Indicator 2 undefined.

Subfield used most often:

$a -- Series statement (R)

$v -- Volume number (R)

```
Example:    490 1# $a Colonial American craftsmen
```

500 General note (R)

Indicators undefined.

Subfield used most often:

$a -- General note (Used when no specialized note field has been defined for the information. Examples: Notes regarding the index; the source of the title; variations in title; descriptions of the nature, form, or scope of the item.)

```
Example:    500 ##  $a Includes index.
```

504 Bibliography, etc. note (R)

Indicators undefined.

Subfield used most often:

$a -- Bibliography, etc. note

```
Example:    504 ## $a Includes bibliographical
                      references.
```

505 Formatted contents note (R)

Indicator 1: Type of contents note

0 -- Complete contents

1 -- Incomplete contents (used with multivolume set when some volumes are not yet published)

2 -- Partial contents

Indicator 2: Level of content designation

-- Basic

Subfield used most often:

$a -- Formatted contents note

```
Example:    505 0# $a Pride and prejudice -- Emma
                      -- Northanger Abbey.
```

520 Summary, etc. note (R)

Indicator 1: Display constant controller

-- Summary
1 -- Review
2 -- Scope and content
3 -- Abstract

Indicator 2 undefined

Subfields used most often

$a -- Summary, abstract, or annotation

$b -- Expansion of summary note

```
Example:    520 ## $a This basic guide to parliamentary
                      procedure tells how to conduct
                      and participate in a meeting
                      properly.
```

600 Subject added entry -- Personal name (R)

Indicator 1: Type of personal name entry element

0 -- Forename
1 -- Surname (this is the most common form)
3 -- Family name

Indicator 2: Subject heading system/thesaurus (identifies the specific list or file which was used)

0 -- Library of Congress Subject Headings
1 -- LC subject headings for children's literature
2 -- Medical Subject Headings
3 -- National Agricultural Library subject authority file
4 -- Source not specified
5 -- Canadian Subject Headings
6 -- Répertoire de vedettes-matière
7 -- Source specified in subfield $2

(Note regarding Sears subject headings: The MARC 21 format does not provide an assigned indicator for Sears subject headings. Therefore, an indicator of 7 is used, and the MARC defined code "sears" is placed in subfield $2.)

Subfields used most often:

$a -- Personal name (surname and forename)
$b -- Numeration
$c -- Titles and other words associated with a name (R)
$q -- Fuller form of name
$d -- Dates associated with a name (generally, year of birth)
$t -- Title of a work
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term (used with 2nd indicator of 7)

```
Example:    600 10 $a Shakespeare, William,
                   $d 1564-1616
                   $x Comedies
                   $x Stage history.
```

```
Example:    600 10 $a Shakespeare, William,
                   $d 1564-1616
                   $x Knowledge
                   $z Rome
                   $v Congresses.
```

Notice that subfields $v, $x, and $z in the 600 field are repeatable. Subfields $v, $x, $y, and $z do not have to be in alphabetical order. They will be in the order prescribed by the instructions given by the subject heading system.

610 Subject added entry -- Corporate name (R)

*Indicator 1:*Type of corporate name entry element
0 -- Inverted name (not used with AACR2)
1 -- Jurisdiction name
2 -- Name in direct order
Indicator 2: Subject heading system/thesaurus.
See indicator 2 under 600
Subfields used most often:
$a -- Corporate name or jurisdiction name as entry element
$b -- Subordinate unit (R)
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term (used with 2nd indicator of 7)

```
Example:    610 10 $a United States.
                   $b Army Air Forces
                   $v Biography.
```

650 Subject added entry -- Topical term (Most subject headings fit here.) (R)

Indicator 1: Level of subject

-- No information provided

Indicator 2: Subject heading system/thesaurus
(identifies the specific list or file which was used)

0 -- Library of Congress Subject Headings
1 -- LC subject headings for children's literature
2 -- Medical Subject Headings
3 -- National Agricultural Library subject authority file
4 -- Source not specified
5 -- Canadian Subject Headings
6 -- Répertoire de vedettes-matière
7 -- Source specified in subfield $2

Note regarding Sears subject headings: The MARC 21 format does not provide an assigned indicator for Sears subject headings. Therefore, an indicator of 7 is used, and the MARC defined code "sears" is placed in subfield $2.)

Subfields used most often:

$a -- Topical term
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term used with 2nd indicator of 7)

```
Example:    650 #0 $a Theater
                   $z United States
                   $v Biography
                   $v Dictionaries.
```

Notice that subfields $v, $x, and $z in the 650 field are repeatable. Subfields $v, $x, $y, and $z do not have to be in alphabetical order. They will be in the order prescribed by the instructions given by the subject heading system.

651 Subject added entry -- Geographic name (R)

Indicator 1: undefined.
Indicator 2: Subject heading system/thesaurus.

See indicator 2 under 600

Subfields used most often:

$a -- Geographic name
$v -- Form subdivision (R)
$x -- General subdivision (R)
$y -- Chronological subdivision (R)
$z -- Geographic subdivision (R)
$2 -- Source of heading or term (used with 2nd indicator of 7)

```
Example:    651 #0 $a United States
                   $x History
                   $v Chronology.
```

Notice that subfields $v, $x, and $z in the 651 field are repeatable. Subfields $v, $x, $y, and $z do not have to be in alphabetical order. They will be in the order prescribed by the instructions given by the subject heading system.

700 Added entry -- Personal name (R)

Indicator 1: Type of personal name entry element

0 -- Forename
1 -- Surname (this is the most common form)
3 -- Family name

Indicator 2: Type of added entry

-- No information provided (most common; co-authors, editors, etc.)
2 -- Analytical entry (The values for Indicator 2 changed in 1994 with Format Integration, and older records may display additional values. An analytical entry involves an author/title of an item contained in a work.)

Subfields used most often:

$a -- Personal name
$b -- Numeration
$c -- Titles and other words associated with a name (R)
$q -- Fuller form of name
$d -- Dates associated with a name (generally, year of birth)
$e -- Relator term (such as ill.) (R)
$4 -- Relator code (R)

```
Example:    700 1# $a Baldridge, Letitia.
```

710 Added entry -- Corporate name (R)

Indicator 1: Type of corporate name entry element

0 -- Inverted name (not used with AACR2)
1 -- Jurisdiction name

2 -- Name in direct order

Indicator 2: Type of added entry.

See Indicator 2 under 700

-- No information provided

2 -- Analytical entry

Subfields used most often:

$a -- Corporate name or jurisdiction name as entry element

$b -- Subordinate unit (R)

```
Example:    710 2# $a Sunburst Communications (Firm)
```

740 Added entry -- Uncontrolled related/analytical title (R)

Indicator 1: Nonfiling characters

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

*Indicator 2:*Type of added entry. See Indicator 2 under 700

-- No information provided

2 -- Analytical entry

(This field was redefined in 1994 with Format Integration. Prior to 1994, the field was also used for variant titles, such as a different wording on a spine title. In records created since Format Integration, those variant tiles appear in a 246 field.)

Subfield used most often:

$a -- Title

```
Example:    740 02 $a Uncle Vanya.
```

800 Series added entry -- Personal name (R)

Indicator 1: Type of personal name entry element

0 -- Forename

1 -- Surname

3 -- Family name

Indicator 2 undefined.

Subfields used most often:

$a -- Personal name

$b -- Numeration

$c -- Titles and other words associated with a name (R)

$q -- Fuller form of name

$d -- Dates associated with a name (generally, year of birth)

$t -- Title of a work (the series)

$v -- Volume number

```
Example:    800 1# $a Fisher, Leonard Everett.
                  $t Colonial American craftsmen.
```

830 Series added entry -- Uniform title (R)

Indicator 1 undefined.

Indicator 2: Nonfiling characters

0-9 -- Number of nonfiling characters present (for initial articles, including spaces)

Subfield used most often:

$a -- Uniform title

$v -- Volume number

```
Example:    830 #0 $a Railroads of America (Macmillan)
```

[Back to Top of Page]

Part VIII:

A List of Other Fields Often Seen in MARC Records

001 Control number
003 Control number identifier
005 Date and time of latest transaction
006 Fixed-length data elements -- additional material characteristics
007 Physical description fixed field
008 Fixed length data elements (See Part X)
022 International Standard Serial Number (ISSN)
037 Source of acquisition
041 Language code
043 Geographic area code
050 Library of Congress call number
060 National Library of Medicine call number
082 Dewey Decimal classification number (the one recommended by the Library of Congress; locally-assigned call numbers may appear elsewhere)
110 Main entry -- Corporate name (less frequent under AACR2 rules)
256 Computer file characteristics

263 Projected publication date
(indicates a CIP -- Cataloging in Publication -- record)
306 Playing time
508 Creation/production credits note
510 Citation/references note (review sources)
511 Participant or performer note
521 Target audience note (first indicator: 0 = reading grade level, 1 = interest age level, 2 = interest grade level, 3 = special audience characteristics, 4 = motivation interest level)
530 Additional physical form available note
538 System details note
586 Awards note
656 Index term -- Occupation
730 Added entry -- Uniform title
852 Location
856 Electronic location and access
9XX Reserved for local use. (They are used by vendors, systems, or individual libraries to exchange additional data)

[Back to Top of Page]

Part IX:

The Leader

There are 24 positions in the Leader, numbered from 00 to 23. For fuller explanation, see the *MARC 21 Format for Bibliographic Data*.

00-04 Record length (calculated by the computer for each record)
05 Record status
- a = increase in encoding level
- c = corrected or revised
- d = deleted
- n = new
- p = increase in encoding from prepublication (previous CIP)

06 Type of record
- a = language material
- c = printed music
- d = manuscript music
- e = cartographic material
- f = manuscript cartographic material
- g = projected medium
- i = nonmusical sound recording
- j = musical sound recording
- k = 2-dimensional nonprojectable graphic
- m = computer file

o = kit
p = mixed materials
r = 3-dimensional artifact or naturally occurring object
t = manuscript language material

07 Bibliographic level
a = monographic component part
b = serial component part
c = collection
d = subunit
i = integrating resource
m = monograph/item
s = serial

08 Type of control
# = no specified type
a = archival

09 Character coding scheme
# = MARC-8
a = UCS/Unicode

10 Indicator count (always "2")

11 Subfield code count (always "2")

12-16 Base address of data (calculated by the computer for each record)

17 Encoding level
# = full level
1 = full level, material not examined
2 = less-than-full level, material not examined
3 = abbreviated level
4 = core level
5 = partial (preliminary) level
7 = minimal level
8 = prepublication level (CIP)
u = unknown
z = not applicable

18 Descriptive cataloging form
# = non-ISBD
a = AACR2
i = ISBD
u = unknown

19 Multipart resource record level
# = Not specified or not applicable
a = Set
b = Part with independent title
c = Part with dependent title

20 Length of the length-of-field portion (always "4")

21 Length of the starting-character-position portion (always "5")

22 Length of the implementation-defined portion (always "0")

23 Undefined (always "0")

[Back to Top of Page]

Part X:

Field 008 for Books

Field 008 is used for Fixed Length Data Elements ("Fixed Field Codes"). There are 40 character positions in field 008, numbered from 00-39. Undefined positions must contain either a blank (#) or a fill character (|). Positions 00-17 and 35-39 are defined the same way for all media.

The information shown here for positions 18-34 applies only to books. For explanation of all the positions below and for positions 18-34 for other media, see the *MARC 21 Format for Bibliographic Data*.

Note that field 008 has no indicators or subfield codes.

00-05 Date entered on file (YYMMDD),
where Y=year, M=month, and D=day

06 Type of date/publication status:
- b = no dates given; B.C. date involved
- e = detailed date
- s = single known date/probable date
- m = multiple dates
- r = reprint/reissue date (Date 1) and original date (Date 2)
- n = dates unknown
- q = questionable date
- t = publication date and copyright date
- | = no attempt to code

07-10 Date 1/beginning date of publication

11-14 Date 2/ending date of publication

Date fields contain the year(s) of publication. The type of date(s) in these elements are specified in fixed field element 06: Type of date/publication status. (For further details, see the field 008 description in the *MARC 21 Format for Bibliographic Data*.)

15-17 Place of publication, production, or execution
For example:
- pk# = Pakistan
- cau = California (US)

(For a full list of codes used in these positions, see the *MARC Code List for Countries*.)

18-21 Illustrations (up to 4 codes):
- # = no illustrations
- a = illustrations
- b = maps
- c = portraits

d = charts
e = plans
f = plates
g = music
h = facsimiles
i = coats of arms
j = genealogical tables
k = forms
l = samples
m = phonodisc, phonowire, etc.
o = photographs
p = illuminations
| = no attempt to code

22 Target audience:
# = unknown or not specified
a = preschool
b = primary
c = pre-adolescent
d = adolescent
e = adult
f = specialized
g = general
j = juvenile
| = no attempt to code

23 Form of item:
# = none of the following
a = microfilm
b = microfiche
c = microopaque
d = large print
f = braille
r = regular print reproduction
s = electronic
| = no attempt to code

24-27 Nature of contents (up to 4):
# = no specified nature of contents
a = abstracts/summaries
b = bibliographies (is one or contains one)
c = catalogs
d = dictionaries
e = encyclopedias
f = handbooks
g = legal articles
i = indexes
j = patent document
k = discographies
l = legislation

m = theses
n = surveys of literature
o = reviews
p = programmed texts
q = filmographies
r = directories
s = statistics
t = technical reports
u = standards/specifications
v = legal cases and notes
w = law reports and digests
z = treaties
| = no attempt to code

28 Government publication:
# = not a government publication
i = international intergovernmental
f = federal/national
a = autonomous or semi-autonomous component
s = state, provincial, territorial, dependent, etc.
m = multistate
c = multilocal
l = local
z = other type of government publication
o = government publication -- level undetermined
u = unknown if item is government publication
| = no attempt to code

29 Conference publication:
0 = not a conference publication
1 = conference publication
| = no attempt to code

30 Festschrift:
0 = not a festschrift
1 = festschrift
| = no attempt to code

31 Index:
0 = no index
1 = index present
| = no attempt to code

32 Undefined (since 1990) (Earlier records may contain the values 0 or 1)
= Undefined
| = no attempt to code

33 Literary form:
0 = not fiction (not further specified)
1 = fiction (not further specified)
c = comic strips
d = dramas
e = essays

f = novels
h = humor, satires, etc.
i = letters
j = short stories
m = mixed forms
p = poetry
s = speeches
u = unknown
| = no attempt to code

34 Biography:
# = no biographical material
a = autobiography
b = individual biography
c = collective biography
d = contains biographical information
| = no attempt to code

35-37 Language:
A three-letter code. For example: eng fre ger spa rus ita

(For a full list of codes used in these positions, see the *MARC Code List for Languages*.)

38 Modified record:
# = not modified
x = missing characters (because of characters unavailable in MARC character set)
s = shortened
d = "dashed-on" information omitted
r = completely romanized/printed cards in script
o = completely romanized/printed cards romanized
| = no attempt to code

39 Cataloging source:
# = national bibliographic agency
c = cooperative cataloging program
d = other sources
u = unknown
| = no attempt to code

[Back to Top of Page]

[Back to Table of Contents] -- [Continue to Part 11]

Library of Congress

Library of Congress Help Desk (10/27/2009)

APPENDIX C

LIBRARY OF CONGRESS | CATALOG

BOOK

Principles of CMOS VLSI design : a systems perspective

Full Record | MARC Tags

Personal name
Weste, Neil H. E.

Main title
Principles of CMOS VLSI design : a systems perspective / Neil H.E. Weste, Kamran Eshraghian.

Edition
2nd ed.

Published/Created
Reading, Mass. : Addison-Wesley Pub. Co., c1993.

Request this Item | LC Find It | Item Availability

LCCN Permalink	https://lccn.loc.gov/92016564
Description	xxii, 713 p., [16] p. of plates : ill. (some col.) ; 24 cm.
ISBN	0201533766
LC classification	TK7874 .W46 1993
Related names	Eshraghian, Kamran.
LC Subjects	Integrated circuits--Very large scale integration--design and construction. Metal oxide semiconductors, Complementary.
Browse by shelf order	TK7874
Notes	Includes bibliographical references and index.
Series	The VLSI systems series
LCCN	92016564
Dewey class no.	621.3/95
Type of material	

Book

Item Availability

CALL NUMBER	TK7874 .W46 1993 CABIN BRANCH Copy 1
Request in	Jefferson or Adams Building Reading Rooms - STORED OFFSITE
Status	c.1 Charged - Due - (Internal Loan)

APPENDIX D

LIBRARY OF CONGRESS | CATALOG

BOOK

Principles of CMOS VLSI design : a systems perspective

Full Record | MARC Tags

000		01212pam a2200301 a 4500
001		1559597
005		20170714131204.0
008		920421s1993 mauaf b 001 0 eng
906	__	|**a** 7 |**b** cbc |**c** orignew |**d** 1 |**e** ocip |**f** 19 |**g** y-gencatlg
925	0_	|**a** acquire |**b** 1 shelf copy |**x** policy default
955	__	|**a** pc01 to ea00 04-21-92; ea10 to SCD 04-21-92; fg11 04-22-92; fl45 04-28-92; CIP ver. pv07 06-16-93
010	__	|**a** 92016564
020	__	|**a** 0201533766
035	__	|**9** (DLC) 92016564
040	__	|**a** DLC |**c** DLC |**d** DLC
050	00	|**a** TK7874 |**b** .W46 1993
082	00	|**a** 621.3/95 |**2** 20
100	1_	|**a** Weste, Neil H. E.
245	10	|**a** Principles of CMOS VLSI design : |**b** a systems perspective / |**c** Neil H.E. Weste, Kamran Eshraghian.
250	__	|**a** 2nd ed.
260	__	|**a** Reading, Mass. : |**b** Addison-Wesley Pub. Co., |**c** c1993.
300	__	|**a** xxii, 713 p., [16] p. of plates : |**b** ill. (some col.) ; |**c** 24 cm.
440	_4	|**a** The VLSI systems series
504	__	|**a** Includes bibliographical references and index.
650	_0	|**a** Integrated circuits |**x** Very large scale integration |**x** design and construction.
650	_0	|**a** Metal oxide semiconductors, Complementary.
700	1_	|**a** Eshraghian, Kamran.
991	__	|**p** 00001496268 |**t** Copy 1 |**b** c-GenColl |**h** TK7874 |**i** .W46 1993 |**w** BOOKS

Request this Item | LC Find It | Item Availability

Item Availability

CALL NUMBER TK7874 .W46 1993 CABIN BRANCH

	Copy 1
Request in	Jefferson or Adams Building Reading Rooms - STORED OFFSITE
Status	c.1 Charged - Due - (Internal Loan)

APPENDIX E

Copyright United States Copyright Office

Help Search History Titles Start Over

Public Catalog

Copyright Catalog (1978 to present)
Search Request: Left Anchored Title = Principles of CMOS VLSI design
Search Results: Displaying 1 of 2 entries

previous next

Labeled View

Principles of CMOS VLSI design : a systems perspective / Neil H. E. Weste,...

Type of Work: Text
Registration Number / Date: TX0003625435 / 1993-08-20
Title: Principles of CMOS VLSI design : a systems perspective / Neil H. E. Weste, Kamran Eshraghian.
Edition: 2nd ed.
Imprint: Reading, MA : Addison-Wesley Pub. Co., c1993.
Description: 713 p.
Copyright Claimant: AT&T Bell Telephone Laboratories, Inc.
Date of Creation: 1992
Date of Publication: 1993-04-26
Variant title: Principles of CMOS VLSI design : a systems perspective
Names: Weste, Neil H. E.
Eshraghian, Kamran
AT&T Bell Telephone Laboratories, Inc.

previous next

Save, Print and Email (Help Page)
Select Download Format Full Record Format for Print/Save
Enter your email address: Email

Help Search History Titles Start Over

APPENDIX F

Simultaneous Driver and Wire Sizing for Performance and Power Optimization

Jason Cong, *Member, IEEE,* and Cheng-Kok Koh

***Abstract*— In this paper, we study the *simultaneous driver and wire sizing (SDWS) problem* under two objective functions: i) delay minimization only, or ii) combined delay and power dissipation minimization. We present general formulations of the SDWS problem under these two objectives based on the distributed Elmore delay model with consideration of both capacitive power dissipation and short-circuit power dissipation. We show several interesting properties of the optimal SDWS solutions under the two objectives, including an important result (Theorem 5) which reveals the relationship between driver sizing and optimal wire sizing. These results lead to polynomial time algorithms for computing the lower and upper bounds of optimal SDWS solutions under the two objectives, and efficient algorithms for computing optimal SDWS solutions under the two objectives. We have implemented these algorithms and compared them with existing design methods for driver sizing only or independent driver and wire sizing. Accurate SPICE simulation shows that our methods reduce the delay by up to 12%–49% and power dissipation by 26%–63% compared with existing design methods.**

I. Introduction

DELAY MINIMIZATION and power dissipation minimization are two important objectives in the design of the high-performance, portable, and wireless computing and communication systems. We believe that both device design (i.e., transistor/cell design) and interconnect design have to be considered and optimized simultaneously in order to achieve these two objectives. This new design methodology is especially important as the VLSI fabrication technology advances to submicron device dimension and gigahertz clock frequency. In this case, interconnect delay becomes an important factor in determining circuit speed, the *distributed* nature of the interconnect structure must be considered, and the capacitive power dissipation on charging/discharging interconnect structures takes a significant portion of overall system power. The objective of this paper is to study the simultaneous driver and wire sizing problem for both delay and power optimization.

In the past, two methods are commonly used to improve the performance of long interconnect lines. One method is driver sizing, which uses a large driver or a series of cascaded drivers of increasing sizes to drive long interconnect lines [2]. Another method is to break long interconnect lines into shorter segments by inserting repeaters. These repeaters can also be sized properly for further reduction in interconnect delay [2]. Both methods are effective for interconnect delay reduction but with substantial increase in power consumption.

Recent studies show that interconnect delay can also be reduced by interconnect topology optimization and wiresizing optimization. A number of interconnect topologies have been proposed for interconnect performance optimization, including bounded-radius bounded-cost trees [7], AHHK trees [1], maximum performance trees [6], *A*-trees [11], low-delay trees [3], [4], and IDW/CFD trees [16]. Wiresizing optimization was first proposed in [11], [12] to further minimize interconnect delay by optimally assigning different wire width to each wire segment in routing trees and substantial delay reduction was achieved for submicron IC design and MCM design. Follow-up work on wiresizing includes the use of convex programming technique [20] and interleave of wiresizing with routing tree construction [15]. Both interconnect topology optimization and wiresizing optimization are effective when resistance ratio, i.e., the driver resistance versus unit wire resistance, is small in the design [11].

Very recently, Cong, Koh, and Leung [8] explored the possibility of both driver sizing and wire sizing. A simple heuristic algorithm was used to size drivers according to a fixed constant ratio, and an *independent* wire sizing optimization is performed for each driver sizing solution. Although encouraging experimental results were reported, it is not difficult to see that this method often produces suboptimal solutions since diver sizing and wire sizing were carried out independently.

In this paper, we study the *simultaneous driver and wire sizing (SDWS) problem* under two objective functions: i) delay minimization only, or ii) combined delay and power dissipation minimization. We present general formulations of the SDWS problem under these two objectives based on the distributed Elmore delay model with consideration of both capacitive power dissipation and short-circuit power dissipation. We show several interesting properties of the optimal SDWS solutions under the two objectives, including an important result (Theorem 5) which reveals the relationship between driver sizing and optimal wire sizing. These results lead to polynomial time algorithms for computing the lower and upper bounds of the optimal SDWS solution under the two objectives, and efficient algorithms for computing optimal SDWS solutions under the two objectives. We have implemented these algorithms and compared them with existing design methods for driver sizing only or independent driver and wire sizing. Accurate SPICE simulation shows that our methods reduce the delay by up to 12%–49% and power dissipation

Manuscript received June 15, 1994; revised August 23, 1994. This work was supported by ARPA/CSTO under Contract J-FB1-93-112, the National Science Foundation Young Investigator Award MIP-9357582, and a Grant from Intel Corporation.

The authors are with the Department of Computer Science, University of California, Los Angeles, CA 90024 USA.

IEEE Log Number 9406370.

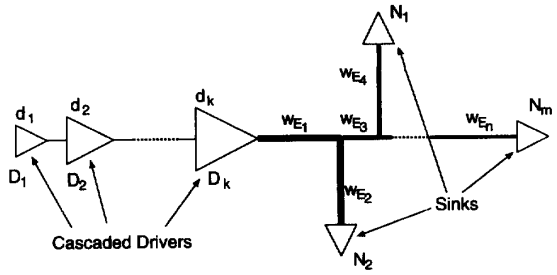


Fig. 1. A k-stage cascaded drivers driving an interconnect tree T with sinks $\{N_1, N_2, \ldots, N_m\}$. w_{E_i} denotes the width of the wire segment E_i, $i = 1 \cdots n$ and d_i denotes the size of driver D_i at ith-stage, $i = 1 \cdots k$.

by 26%–63% compared with existing design methods. To the best of our knowledge, this is the first work which presents in-depth study of the SDWS problem for both delay and power optimization.

The remainder of this paper is organized as follows. In Section II, we present the general formulation of the simultaneous driver sizing and wiresizing problems under the delay minimization objective and the combined delay and power minimization objective. In Section III, we present results on optimal SDWS solutions for delay minimization. In Section IV, we present results on optimal SDWS solutions under the combined delay and power minimization objective. Section V shows the experimental results obtained by our SDWS algorithms and the comparative study with other existing methods. Section VI concludes the paper with discussions of future work. An extended abstract of this paper was presented in ICCAD'94 [10].

II. Problem Formulation

A. Performance Optimization

Assume that we are given a routing tree T implementating a signal net which consists of a source N_+, and a set of m sinks $\{N_1, N_2, \ldots, N_m\}$. A node in T refers to the source, or a sink, or a Steiner node, and a segment connects two nodes in T. Assume that $\{E_1, E_2, \ldots, E_n\}$ is the set of segments forming the tree T, where n is the total number of segments in the tree. Each wire segment has a set of discrete choices of wire widths $\{W_1, W_2, \ldots, W_r\}$ $(W_1 < W_2 < \cdots < W_r)$. We use w_{E_i} to denote the width of the wire segment E_i, $i = 1 \cdots n$.

Furthermore, we assume that the signal net is driven by a chain of cascaded drivers of k stages at the source as shown in Fig. 1. We use $\mathcal{D} = \{d_1, d_2, \ldots, d_k\}$ to denote a driver sizing solution, where d_i denotes the size of driver D_i at i-th stage $(i = 1 \cdots k)$. We assume that driver D_1 is of minimum size, i.e., $d_1 = 1$ (after normalization). Given the above definitions, the problem of simultaneous driver and wire sizing (SDWS) for performance optimization can be defined as follows:

Formulation 1: Given a routing tree T, the SDWS problem for delay minimization (SDWS-D) is to determine the number of stages k, a driver sizing solution $\mathcal{D}$, and a wiresizing solution $\mathcal{W}$ on T, such that the performance measure t $(T, k, \mathcal{D}, \mathcal{W})$ is optimized.

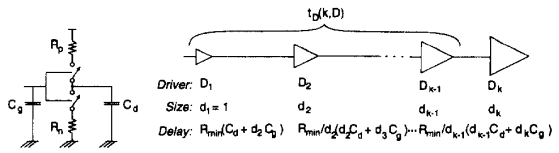


Fig. 2. (a) A switch-level *RC* model of a driver (b) Inter-stage delay of a k-stage cascaded drivers until the gate of the k-th driver.

If we fix the number of stages k, a restricted version of the SDWS-D problem called the k-SDWS-D problem can be defined as follows: Given a routing tree T and a chain of k drivers, the k-SDWS-D problem is to determine the optimal driver and wire sizing solution $\mathcal{D}$ and $\mathcal{W}$, such that the performance measure t $(T, k, \mathcal{D}, \mathcal{W})$ is optimized.

The performance measure t $(T, k, \mathcal{D}, \mathcal{W})$ approximates the delay of the signal net from the source to one or several critical sinks, and it can be decomposed as follows:

$$t(T, k, \mathcal{D}, \mathcal{W}) = t_D(k, \mathcal{D}) + t_I(T, k, \mathcal{D}, \mathcal{W}) \qquad (1)$$

where the first term $t_D(k, \mathcal{D})$ measures the delay due to the drivers and the second term $t_I(T, k, \mathcal{D}, \mathcal{W})$ measures the interconnect delay. We estimate $t_D(k, \mathcal{D})$ based on a switch-level *RC* model of drivers. The interconnect delay $t_I(T, k, \mathcal{D}, \mathcal{W})$ is measured by the distributed Elmore delay model [14].

RC Delay Model for Drivers: We estimate the delay of a driver based on a switch-level *RC* model of driver. Fig. 2(a) shows a minimum-size inverter (driver) with a p-transistor resistance R_p and a n-transistor resistance R_n. We assume that $R_p = R_n = R_{\min}$. Let C_g denote the gate capacitance and C_d denote the capacitance due to the source and drain diffusion of the minimum-size driver.

Fig. 2(b) illustrates the delay due to a sequence of k cascaded drivers $\mathcal{D}$. Ignoring the interconnect resistance and capacitance between drivers (since cascaded drivers are placed closed to each other), the delay from driver D_i to D_{i+1} $(1 \le i \le k-1)$ is the product of the resistance of D_i and its capacitive load:

$$\frac{R_{\min}}{d_i} \cdot (C_d \cdot d_i + C_g \cdot d_{i+1}) = R_{\min} \cdot C_d + R_{\min} \cdot C_g \cdot \frac{d_{i+1}}{d_i} \qquad (2)$$

The total delay up to gate of the last driver $t_D(k, \mathcal{D})$ (excluding the last driver) can be expressed as follows:

$$t_D(k, \mathcal{D}) = (k-1) \cdot R_{\min} \cdot C_d + \sum_{i=1}^{k-1} \frac{R_{\min} \cdot C_g \cdot d_{i+1}}{d_i}$$

$$= \mathcal{J}_1 + \mathcal{J}_2 \cdot \sum_{i=1}^{k-1} \frac{d_{i+1}}{d_i} \qquad (3)$$

where $\mathcal{J}_1 = (k-1) \cdot R_{\min} \cdot C_d$, and $\mathcal{J}_2 = R_{\min} \cdot C_g$. Notice that delay through the k-th driver will be counted as part of interconnect delay in Section II-A-3.

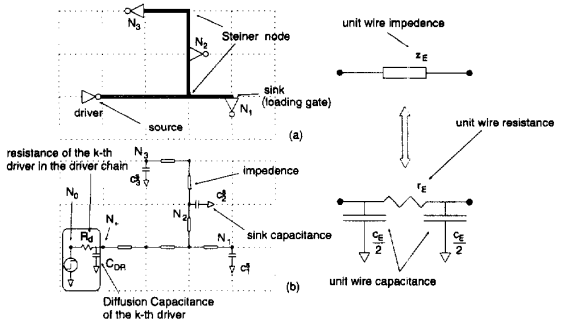


Fig. 3. A uniform grid structure for the distributed *RC* interconnect model. (a) The layout of an interconnect tree $T$with three sinks N_1, N_2 and N_3. (b) The corresponding distributed *RC* interconnect model of T. Each grid edge E in T connecting two adjacent nodes is modeled as a π-type *RC* circuit containing a resistor of r_E and two capacitors of $\frac{c_E}{2}$ each. Each sink has an extra loading capacitance.

Distributed Elmore Delay Model for Interconnect: We use the distributed Elmore delay model [14] for interconnect delay measure. The formulations used in this section are based on those in [12]. In order to model a routing tree as a distributed *RC* circuit accurately, a uniform grid structure is superimposed on the routing plane, and each wire segment in the routing plane is divided into a sequence of wires of unit length as shown in Fig. 3. In this case, T consists of a set of unit-length wire segments, each may have a different width[1].

For each edge E in the tree T, we use a π-type *RC* circuit to model E, where r_E and c_E are interconnect resistance and capacitance, respectively. Given an edge E, we use w_E to denote the width of the grid edge E. Assume that a unit-width unti-grid-length wire has wire resistance r_0, wire area capacitance c_0 and wire fringing capacitance c_1, then r_E and c_E for any grid edge E can be defined as follows:

$$r_E = \frac{r_0}{w_E}$$
$$c_E = c_0 \cdot w_E + c_1 \tag{4}$$

Note that previous formulations in [11] and [12] did not consider fringing effect. Also, c_1 can be used to represent mutual capacitance between adjacent wires in the same layer or adjacent layers, when it is known.

We use c_u^s to denote the capacitance at sink u. To correctly model the driver resistance, we introduce an additional node N_0 and connect N_0 to N_+ via an additional segment with resistance $R_d = R_{\min}/d_k$ and capacitance $C_{\rm DR} = C_d \cdot d_k$ (the resistance and capacitance of the last driver in the cascaded driver chain) in the later discussions.

Given a grid edge E, we use $\mathrm{Des}(E)$ to denote the set of grid edges in the subtree rooted at E (excluding E), and $\mathrm{Ans}(E)$ to denote the set of grid edges $\{E' \mid E \in \mathrm{Des}(E')\}$ (again, excluding E). That is, $\mathrm{Des}(E)$ is the set of "descendant" grid edges of E, and $\mathrm{Ans}(E)$ is the set of "ancestor" grid edges of E. Also, we use $\mathrm{sink}(E)$ to denote the set of sinks in the subtree rooted at E, and C_E to denote the *total* capacitance in the subtree rooted at E (including both the wire capacitances and the sink capacitances):

$$C_E = \sum_{u \in \mathrm{sink}(E)} c_u^s + c_0 \cdot \sum_{E' \in \mathrm{Des}(E)} w_{E'} + \sum_{E' \in \mathrm{Des}(E)} c_1 \tag{5}$$

Furthermore, we use $\mathrm{sink}(T)$ to denote the set of sinks in T and $P(u, v)$ to denote the unique path from u to v for any grid points u, v in the routing tree.

We use the Elmore delay model [14] as the objective function for delay optimization. Given a distributed *RC* circuit tree T, the signal delay at a particular node N_i, denoted as $t(N_i)$, is computed as follows:

$$t(N_i) = \sum_{E \in P(N_0, N_i)} r_E \cdot \left(\frac{c_E}{2} + C_E\right) \tag{6}$$

where the summation is taken over all the grid edges on the path from the driver N_0 to the node N_i.

Single Critical Sink Formulation: We first study the case where there is only one critical sink N_i in the net. Let R_d and $C_{\rm DR}$ be the resistance and diffusion capacitance of the last driver in the driver chain, respectively, i.e., $R_d = R_{\min}/d_k$, $C_{\rm DR} = C_d \cdot d_k$. Also, we use $C_{\rm IL}(T, \mathcal{W})$ to denote total capacitance of interconnect tree T (including the loading capacitance at the sinks). We define

$$\begin{aligned} C_{\rm IL}(T, \mathcal{W}) &= \sum_{u \in \mathrm{sink}(T)} c_u^s + c_0 \cdot \sum_{E \in T} w_E + \sum_{E \in T} c_1 \\ &= \sum_{u \in \mathrm{sink}(T)} c_u^s + c_0 \cdot \sum_{E \in T} w_E + n \cdot c_1 \end{aligned} \tag{7}$$

which is the total capacitance that the driver chain will drive. According to (6) and following a similar derivation[2] in [11] and [12], the signal delay $t_i(T, R_d, \mathcal{W})$ at N_i under a given wiresizing solution $\mathcal{W}$ is:

$$\begin{aligned} & t_i(T, R_d, \mathcal{W}) \\ &= R_d \cdot (C_{\rm DR} + C_{\rm IL}(T, \mathcal{W})) \\ &\quad + \sum_{E \in P(N_+, N_i)} r_E \cdot \left(\frac{c_E}{2} + C_E\right) \\ &= R_d \cdot C_{\rm DR} + R_d \cdot \sum_{u \in \mathrm{sink}(T)} c_u^s \\ &\quad + R_d \cdot c_0 \cdot \sum_{E \in T} w_E + R_d \cdot n \cdot c_1 \\ &\quad + \sum_{E \in P(N_+, N_i)} \frac{r_0 \cdot c_0}{2} + \frac{r_0 \cdot c_1}{2} \cdot \sum_{E \in P(N_+, N_i)} \frac{1}{w_E} \\ &\quad + r_0 \cdot c_0 \cdot \sum_{E \in P(N_+, N_i)} \sum_{E' \in \mathrm{Des}(E)} \frac{w_{E'}}{w_E} \\ &\quad + r_0 \cdot c_1 \cdot \sum_{E \in P(N_+, N_i)} \sum_{E' \in \mathrm{Des}(E)} \frac{1}{w_E} \\ &\quad + r_0 \cdot \sum_{E \in P(N_+, N_i)} \sum_{v \in \mathrm{sink}(E)} c_v^s \cdot \frac{1}{w_E} \end{aligned} \tag{8}$$

[1] In [11] and [12], a segment in T is defined to be an edge in the rectilinear Steiner tree. In that case, segments in T are of different lengths, and the wire width is uniform within each segment.

[2] In this paper, detailed derivation of (8), (12), (23), and (24) are omitted. The details can be found in the technical report [9].

Let

$$\begin{aligned}
\mathcal{K}_0^i &= R_d \cdot C_{\mathrm{DR}} + R_d \cdot \sum_{u \in \mathrm{sink}(T)} c_u^s + R_d \cdot n \cdot c_1 \\
&\quad + \sum_{E \in P(N_+, N_i)} \frac{r_0 \cdot c_0}{2} \\
\mathcal{K}_1 &= R_d \cdot c_0 \\
\mathcal{K}_2 &= r_0 \cdot c_0 \\
\mathcal{K}_3 &= r_0 \cdot c_1 \\
\mathcal{K}_4 &= r_0 \\
\mathcal{K}_5 &= \frac{r_0 \cdot c_1}{2}
\end{aligned} \tag{9}$$

Note that given a fixed $R_d, \mathcal{K}_0^i$ is a constant with respect to the timing-critical sink N_i, and $\mathcal{K}_1$ to $\mathcal{K}_5$ are all constants depending only on the technology. Also, we define the functions $f_i(E, E'), g_i(E)$ and $h_i(E)$ as follows:

$$f_i(E, E') = \begin{cases} 1 & \text{if } E \in P(N_+, N_i) \text{ and } E' \in \mathrm{Des}(E) \\ 0 & \text{otherwise} \end{cases}$$

$$g_i(E) = \begin{cases} \sum_{v \in \mathrm{sink}(E)} c_v^s & \text{if } E \in P(N_+, N_i) \\ 0 & \text{otherwise} \end{cases}$$

$$h_i(E) = \begin{cases} 1 & \text{if } E \in P(N_+, N_i) \\ 0 & \text{otherwise} \end{cases}$$

Rearranging, we can rewrite (8) as follows:

$$\begin{aligned}
t_i(T, R_d, \mathcal{W}) &= \mathcal{K}_0^i + \mathcal{K}_1 \cdot \sum_{E \in T} w_E \\
&\quad + \mathcal{K}_2 \cdot \sum_{E, E' \in T} f_i(E, E') \cdot \frac{w_{E'}}{w_E} \\
&\quad + \mathcal{K}_3 \cdot \sum_{E, E' \in T} f_i(E, E') \cdot \frac{1}{w_E} \\
&\quad + \mathcal{K}_4 \cdot \sum_{E} g_i(E) \cdot \frac{1}{w_E} \\
&\quad + \mathcal{K}_5 \cdot \sum_{E \in T} h_i(E) \cdot \frac{1}{w_E}
\end{aligned} \tag{10}$$

Multiple Critical Sinks Formulation: Let $\mathrm{sink}(T)$ denote the set of sinks in T. When there are several critical sinks of different priorities in the routing tree, the previous formulation can be generalized as follows:

$$t_i(T, R_d, \mathcal{W}) = \sum_{N_i \in \mathrm{sink}(T)} \lambda_i \cdot t_i(T, R_d, \mathcal{W}) \tag{11}$$

where λ_i is the weight of the delay penalty to sink N_i [3], [12]. The larger λ_i is, the more critical sink N_i is. We normalize λ_i's such that $\sum_{N_i \in \mathrm{sink}(T)} \lambda_i = 1$. We can rewrite (11) as follows:

$$\begin{aligned}
t(T, R_d, \mathcal{W}) &= \mathcal{K}_0 + \mathcal{K}_1 \cdot \sum_{E \in T} w_E \\
&\quad + \mathcal{K}_2 \cdot \sum_{E, E' \in T} F(E, E') \cdot \frac{w_{E'}}{w_E} \\
&\quad + \mathcal{K}_3 \cdot \sum_{E, E' \in T} F(E, E') \cdot \frac{1}{w_E} \\
&\quad + \mathcal{K}_4 \cdot \sum_{E \in T} G(E) \cdot \frac{1}{w_E} \\
&\quad + \mathcal{K}_5 \cdot \sum_{E \in T} H(E) \cdot \frac{1}{w_E}
\end{aligned} \tag{12}$$

where

$$\begin{aligned}
\mathcal{K}_0 &= \sum_{N_i \in \mathrm{sink}(T)} \lambda_i \cdot \mathcal{K}_0^i \\
F(E, E') &= \sum_{N_i \in \mathrm{sink}(T)} \lambda_i \cdot f_i(E, E') \\
G(E) &= \sum_{N_i \in \mathrm{sink}(T)} \lambda_i \cdot g_i(E) \\
H(E) &= \sum_{N_i \in \mathrm{sink}(T)} \lambda_i \cdot h_i(E)
\end{aligned} \tag{13}$$

Using a similar argument as in [13], we have

$$\begin{aligned}
F(E_1, E_2) &\geq F(E_1, E_2') && \text{if } E_2 \in \mathrm{Des}(E_2') \\
F(E_1, E_2) &\geq F(E_1', E_2) && \text{if } E_1 \in \mathrm{Ans}(E_1') \\
G(E_1) &\geq G(E_1') && \text{if } E_1 \in \mathrm{Ans}(E_1') \\
H(E_1) &\geq H(E_1') && \text{if } E_1 \in \mathrm{Ans}(E_1') \\
H(E_1) &\geq F(E_1, E_2) && \text{if } E_2 \in \mathrm{Des}(E_1)
\end{aligned} \tag{14}$$

Performance Optimization Objective of the SDWS Problem: Let $t_I(T, k, \mathcal{D}, \mathcal{W}) = t(T, R_d, \mathcal{W})$. Hence, the performance measure on both driver and interconnect delay $t(T, k, \mathcal{D}, \mathcal{W})$ in (1) can be written as

$$\begin{aligned}
&t(T, k, \mathcal{D}, \mathcal{W}) \\
&\quad = t_D(k, \mathcal{D}) + t_I(T, k, \mathcal{D}, \mathcal{W}) \\
&\quad = \left\{ \mathcal{J}_1 + \mathcal{J}_2 \cdot \sum_{i=1}^{k-1} \frac{d_{i+1}}{d_i} \right\} + \left\{ \mathcal{K}_0 + \mathcal{K}_1 \cdot \sum_{E \in T} w_E \right. \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E, E' \in T} F(E, E') \cdot \frac{w_{E'}}{w_E} + \mathcal{K}_3 \cdot \sum_{E, E' \in T} F(E, E') \\
&\qquad \left. \cdot \frac{1}{w_E} + \mathcal{K}_4 \cdot \sum_{E \in T} G(E) \cdot \frac{1}{w_E} + \mathcal{K}_5 \cdot \sum_{E \in T} H(E) \cdot \frac{1}{w_E} \right\}
\end{aligned} \tag{15}$$

where $\mathcal{K}_0, \mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3, \mathcal{K}_4$, and $\mathcal{K}_5$ are constants defined in (9) and (13).

B. Performance and Power Optimization

Driver and wire sizing are effective approaches to reduce interconnect delay. However, larger driver size and additional routing capacitance also increase the power dissipation. In practice, high-speed circuit design requires a careful trade-off between performance and power. We define the SDWS problem for both delay and power optimization (SDWS-DP) as follows:

Formulation 2: Given a routing tree T, the SDWS problem for both delay and power optimization (SDWS-DP) is to determine the number of stages k, a driver sizing solution $\mathcal{D}$, and a wiresizing solution $\mathcal{W}$ on T, such that the objective function $\mathrm{obj}(T, k, \mathcal{D}, \mathcal{W})$ defined below is minimized:

$$\begin{aligned}
\mathrm{obj}(T, k, \mathcal{D}, \mathcal{W}) &= \alpha \cdot \mathrm{Power}(T, k, \mathcal{D}, \mathcal{W}) \\
&\quad + \gamma \cdot t(T, k, \mathcal{D}, \mathcal{W})
\end{aligned} \tag{16}$$

where $t(T,k,\mathcal{D},\mathcal{W})$ is the performance measure, Power $(T,k,\mathcal{D},\mathcal{W})$ is the power dissipation, and α and γ are adjustable non-negative parameters controlling the trade-off between performance and power dissipation.

Parameters α and γ can either be given by designer or higher level synthesis program. Similarly, if we fix the number of stages k, we can define the *k-SDWS-DP* problem as follows: Given a routing tree T and a chain of k drivers, determine the optimal driver and wire sizing solution $\mathcal{D}$ and $\mathcal{W}$, such that the combined objective function obj$(T,k,\mathcal{D},\mathcal{W})$ is minimized.

Power Dissipation Formulation: There are two components that determine the amount of power dissipated in a CMOS circuit, namely *static* dissipation due to leakage current, and *dynamic* dissipation due to switching transient current (short-circuit dissipation) and charging and discharging of load capacitances (capacitive dissipation) [21]. We consider only dynamic dissipation in our formulation since static dissipation is usually 2 to 3 order of magnitude smaller. Given a loading capacitance C_L, we can write the capacitive and short-circuit dissipation of a single driver as follows [21]:

$$\text{Power}_{\text{cap}} = f \cdot C_L \cdot V_{dd}^2 \tag{17}$$

$$\text{Power}_{\text{sc}} = f \cdot \frac{\beta}{12} \cdot (V_{dd} - 2V_t)^3 \cdot t_{\text{rf}} \tag{18}$$

where f is the switching frequency of the input signal, β is the MOS transistor gain factor, V_t is the threshold voltage, and t_{rf} is the rise and fall time of the input signal which are assumed to be equal.

Consider a chain of k drivers. For driver $D_i (i = 1 \cdots k-1)$, the capacitive load of the driver is assumed to be the sum of its diffusion capacitance $C_d \cdot d_i$ and the gate capacitance $C_g \cdot d_{i+1}$ of the next driver D_{i+1}. The last driver D_k has a capacitive load of $C_d \cdot d_k + C_{\text{IL}}(T,\mathcal{W})$. Hence the capacitive power of the cascaded drivers is

$$\begin{aligned}
&\text{Power}_{\text{cap}}(k,\mathcal{D},C_{\text{IL}}(T,\mathcal{W})) \\
&\quad = f \cdot V_{dd}^2 \cdot \left(C_d + \sum_{i=2}^{k-1} (d_i \cdot C_d + d_{i+1} \cdot C_g) \right. \\
&\qquad \left. + \; d_k \cdot C_d + C_{\text{IL}}(T,\mathcal{W}) \right) \\
&\quad = \mathcal{L}_0 + \mathcal{L}_1 \cdot \left(\sum_{i=2}^{k} d_i + \frac{C_{\text{IL}}(T,\mathcal{W})}{C_g + C_d} \right)
\end{aligned} \tag{19}$$

where $\mathcal{L}_0 = f \cdot V_{\text{dd}}^2 \cdot C_d$ and $\mathcal{L}_1 = f \cdot V_{\text{dd}}^2 \cdot (C_g + C_d)$.

Let $\beta_{\min}$ be the gain factor of a minimum-size driver. The gain factor of a driver of size d can be defined as $d \cdot \beta_{\min}$. Hence, we can write the short-circuit power of the cascaded drivers as follows [21]:

$$\begin{aligned}
\text{Power}_{\text{sc}}(k,\mathcal{D},C_{\text{IL}}(T,\mathcal{W})) &= f \cdot \sum_{i=1}^{k} \frac{\beta_{\min} \cdot d_i}{12} \cdot (V_{dd} - 2V_t)^3 \cdot t_{\text{rf}} \\
&= \mathcal{L}_2 + \mathcal{L}_2 \cdot \sum_{i=1}^{k} d_i
\end{aligned} \tag{20}$$

where $\mathcal{L}_2 = f \cdot \frac{\beta_{\min}}{12} \cdot (V_{dd} - V_t)^3 \cdot t_{\text{rf}}$. Note that one limitation of this simplified formulation is that it assumes t_{rf} to be constant for all drivers along the chain. In reality, the rise and fall times of the transition at the input to the gate depend on the driving capability of the previous stage driver, the loading capacitance of the gate concerned and the interconnect between the drivers.

Summing up the capacitive and short-circuit power, the dynamic power dissipation of the circuit can be written as:

$$\begin{aligned}
\text{Power}(T,k,\mathcal{D},\mathcal{W}) &= \text{Power}_{\text{cap}}(k,\mathcal{D},C_{\text{IL}}(T,\mathcal{W})) \\
&\quad + \text{Power}_{\text{sc}}(k,\mathcal{D},C_{\text{IL}}(T,\mathcal{W})) \\
&= \left\{ \mathcal{L}_0 + \mathcal{L}_1 \cdot \left(\sum_{i=2}^{k} d_i + \frac{C_{\text{IL}}(T,\mathcal{W})}{C_g + C_d} \right) \right\} \\
&\quad + \left\{ \mathcal{L}_2 + \mathcal{L}_2 \cdot \sum_{i=2}^{k} d_i \right\} \\
&= \mathcal{L}_0 + \mathcal{L}_1 \cdot \frac{C_{\text{IL}}(T,\mathcal{W})}{C_g + C_d} + \mathcal{L}_2 \\
&\quad + (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2}^{k} d_i
\end{aligned} \tag{21}$$

Trade-Off Between Performance and Power: We can now write the trade-off between performance and power in (16) as follows:

$$\begin{aligned}
&\text{obj}(T,k,\mathcal{D},\mathcal{W}) \\
&\quad = \alpha \cdot \text{Power}(T,k,\mathcal{D},\mathcal{W})) + \gamma \cdot t(T,k,\mathcal{D},\mathcal{W}) \\
&\quad = \left\{ \alpha \cdot \left(\mathcal{L}_0 + \mathcal{L}_1 \cdot \frac{C_{\text{IL}}(T,\mathcal{W})}{C_g + C_d} \right. \right. \\
&\qquad \left. \left. + \mathcal{L}_2 + (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2}^{k} d_i \right) \right\} \\
&\quad + \left\{ \gamma \cdot \left(\mathcal{J}_1 + (\mathcal{J}_2 \cdot \sum_{i=1}^{k-1} \frac{d_{i+1}}{d_i} + \mathcal{K}_0 + \mathcal{K}_1 \cdot \sum_{E \in T} w_E \right. \right. \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E,E' \in T} F(E,E') \cdot \frac{w_{E'}}{w_E} \\
&\qquad + \mathcal{K}_3 \cdot \sum_{E,E' \in T} F(E,E') \cdot \frac{1}{w_E} \\
&\qquad \left. \left. + \mathcal{K}_4 \cdot \sum_{E \in T} G(E) \cdot \frac{1}{w_E} + \mathcal{K}_5 \cdot \sum_{E \in T} H(E) \cdot \frac{1}{w_E} \right) \right\}
\end{aligned} \tag{22}$$

There are two terms that links the driver chain and the interconnect in both the power formula and delay formula, namely the last driver d_k and the capacitive load $C_{\text{IL}}(T,\mathcal{W})$. Suppose a wire-width assignment $\mathcal{W}$ is given for a routing tree T, the capacitive load $C_{\text{IL}}(T,\mathcal{W})$ can be computed. Eliminating the constant terms, the drivers are sized to minimize the following:

$$\begin{aligned}
\text{obj}_{T,k,\mathcal{W}}(\mathcal{D}) &= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2}^{k} d_i \\
&\quad + \gamma \cdot \mathcal{J}_2 \cdot \left(\frac{C_{\text{IL}}(T,\mathcal{W})}{d_k \cdot C_g} + \sum_{i=1}^{k-1} \frac{d_{i+1}}{d_i} \right)
\end{aligned} \tag{23}$$

Given a driver size assignment, the wires are sized to minimize the following tradeoff between routing area and interconnect

delay after eliminating the constant terms:

$$\begin{aligned}\text{obj}_{T,k,\mathcal{D}}(\mathcal{W}) = &\left(\alpha \cdot \mathcal{L}_1 \cdot \frac{1}{C_g + C_d} \cdot c_0 + \gamma \cdot \mathcal{K}_1\right) \cdot \sum_{E \in T} w_E \\ &+ \gamma \cdot \mathcal{K}_2 \cdot \sum_{E, E' \in T} F(E, E') \cdot \frac{w_{E'}}{w_E} \\ &+ \gamma \cdot \mathcal{K}_3 \cdot \sum_{E, E' \in T} F(E, E') \cdot \frac{1}{w_E} \\ &+ \gamma \cdot \mathcal{K}_4 \cdot \sum_{E \in T} G(E) \cdot \frac{1}{w_E} \\ &+ \gamma \cdot \mathcal{K}_5 \cdot \sum_{E \in T} H(E) \cdot \frac{1}{w_E} \qquad (24)\end{aligned}$$

The formulation is similar to (12) except with difference in constant coefficients, which implies that the wiresizing results for delay minimization can also be applied to simultaneous power and performance optimization.

Since $\text{Power}(T, k, \mathcal{D}, \mathcal{W})$ and $t(T, k, \mathcal{D}, \mathcal{W})$ are usually of different orders of magnitude. The choice of α and γ in (22) might be difficult. We fine in our study that it is convenient to optimize the following objective:

$$\begin{aligned}\text{obj}_\alpha(T, k, \mathcal{D}\mathcal{W}) = &\alpha \cdot \frac{\text{Power}(T, k, \mathcal{D}, \mathcal{W})}{\text{Power}_{\min}(T)} \\ &+ (1 - \alpha) \cdot \frac{t(T, k, \mathcal{D}, \mathcal{W})}{t_{\min}(T, \mathcal{D}, \mathcal{W})} \qquad (25)\end{aligned}$$

where $\text{Power}_{\min}(T)$ is the minimum power required for driving an interconnect tree T and $t_{\min}(T, \mathcal{D}, \mathcal{W})$ is the minimum delay achievable by a chain of drivers driving an interconnect tree T.

Note that given a SDWS-DP problem, $\text{Power}_{\min}(T)$ can be computed easily by assuming a single driver driving an interconnect tree where all interconnect grid edges are assigned with minimum wire width. On the other hand, $t_{\min}(T, \mathcal{D}, \mathcal{W})$ can be computed using the algorithms to be presented in Section III. Therefore, they can be treated as constants.

Let k^*_α, $\mathcal{D}^*_\alpha$ and $\mathcal{W}^*_\alpha$ denote the optimal stage number and the optimal driver and sizing solution, respectively for a given α. Note that as α increases from 0 to 1, $\text{Power}(T, k^*_\alpha, \mathcal{D}^*_\alpha, \mathcal{W}^*_\alpha)$ will decrease monotonically and $t(T, k^*_\alpha, \mathcal{D}^*_\alpha, \mathcal{W}^*_\alpha)$ will increase monotonically[3]. Therefore, we can easily adjust parameter α between 0 and 1 to achieve smooth trade-off between performance and power.

III. Simultaneous Driver and Wire Sizing for Performance Optimization (SDWS-D Problem)

A. Properties of Optimal SDWS-D Solutions

We consider the simultaneous driver and wire sizing problem for performance optimization which minimizes the performance measure $t(T, k, \mathcal{D}, \mathcal{W})$ specified by (15). The lemmas and theorems in this subsection are used to prove the correctness of the bound computation algorithm and optimality of the optimal algorithm to the SDWS-P problem. If the reader is only interested to learn the algorithms, one may proceed directly to Section III-B.

Properties of Optimal Driver Sizing Solutions for Performance Optimization:

Theorem 1: [17], [21] Given the loading capacitance C_{IL}, a chain of k drivers and the minimum gate capacitance C_g, the optimal stage ratio is $s = \left(\frac{C_{\text{IL}}}{C_g}\right)^{1/k}$. □

Note that Theorem 1 assumes that the interconnect capacitances between drivers are negligible. We also assume that the loading capacitance C_{IL} is larger than the minimum gate capacitance C_g. Otherwise, a single minimum size driver is enough to drive the load.

Properties of Optimal Wire Sizing Solutions for Performance Optimization: The results in [12] showed several interesting properties of optimal wiresizing solutions for performance optimization, including the *separability, the monotone property*, and the *dominance property*. But they did not model the driver capacitance and interconnect fringing capacitance. Our study shows that these results still hold after we model the driver capacitance and fringing capacitance in the interconnect delay formulation specified by (12). Proofs of these theorems can be found in [9]. They are similar to the proofs of these properties in [13], except that our proofs can handle the driver capacitance and fringing capacitance of the interconnects.

Theorem 2 (separability): If the wire width assignment of a path P originated from the source is given, the optimal wire width assignment for each subtree branching off from P can be carried out independently. □

Definition 1: Given a routing tree T, a wiresizing solution $\mathcal{W}$ on T is a monotone assignment if $w_E \geq w_{E'}$ for any pair of segments E and E' such that $E \in \text{Ans}(E')$.

Theorem 3 (Monotone Property): For any given tree T, there exists a monotone optimal wire width assignment $\mathcal{W}^*$. □

Definition 2: Given two wire width assignments $\mathcal{W}$ and $\mathcal{W}'$, $\mathcal{W}$ dominates $\mathcal{W}'$ if for any segment E, the width assignment of E in $\mathcal{W}$ is greater than or equal to that of E in $\mathcal{W}'$.

Definition 3: Given a routing tree T, a wire width assignment $\mathcal{W}$ of T, and any particular segment $E \in T$, a local refinement on E is the operation to optimize the width of E based on the objective function in (12), without changing the widths of other segments specified by $\mathcal{W}$.

Theorem 4 (Dominance Property): Given a routing tree T, let $\mathcal{W}^*$ be an optimal wire width assignment. If a wire width assignment $\mathcal{W}$ dominates $\mathcal{W}^*$, then a local refinement of $\mathcal{W}$ on any edge E is T still dominates $\mathcal{W}^*$. Similarly, if a wire width assignment $\mathcal{W}$ is dominated by $\mathcal{W}^*$, then a local refinement of $\mathcal{W}$ on any edge E in T is dominated by $\mathcal{W}^*$. □

Relation Between Driver Sizing and Optimal Wire Sizing for Performance Optimization: Given a routing tree T with one or more critical sinks. Let R_d be the resistance of the last driver driving the routing tree and $\mathcal{W}^*$ be the corresponding optimal wire width assignment. Consider another chain of cascaded drivers such that R'_d is the resistance of the last driver and $\mathcal{W}'^*$ is the optimal wire width assignment. We shall show the

[3] Otherwise, we can swap the optional SDWS-DP solutions for α and $\alpha'(\neq \alpha)$ and show that either $\text{obj}_\alpha(T, k^*_{\alpha'}, \mathcal{D}^*_{\alpha'}, \mathcal{W}^*_{\alpha'}) < \text{obj}_\alpha(T, k^*_\alpha, \mathcal{D}^*_\alpha, \mathcal{W}^*_\alpha)$ or $\text{obj}_{\alpha'}(T, k^*_\alpha, \mathcal{D}^*_\alpha, \mathcal{W}^*_\alpha) < \text{obj}_{\alpha'}(T, k^*_{\alpha'}, \mathcal{D}^*_{\alpha'}, \mathcal{W}^*_{\alpha'})$ to obtain a contradiction.

following result that $R_d < R'_d$ implies that $\mathcal{W}^*$ dominates $\mathcal{W}'^*$ (Theorem 5):

Lemma 1: If $R_d < R'_d$, then area $(\mathcal{W}^*) \geq$ area$(\mathcal{W}'^*)$, where area $(\mathcal{W})$ denotes the total wire width used in wire width assignment $\mathcal{W}$.

Proof: Let $\Delta t(T, R_d, \mathcal{W} \rightarrow \mathcal{W}')$ denote the difference $t(T, R_d, \mathcal{W}') - t(T, R_d, \mathcal{W})$. Clearly, $\Delta t(T, R_d, \mathcal{W}^* \rightarrow \mathcal{W}'^*) \leq 0$ and $\Delta t(T, R'_d, \mathcal{W}'^* \rightarrow \mathcal{W}^*) \leq 0$ since $\mathcal{W}^*$ and $\mathcal{W}'^*$ are the optimal wiresizing solutions for R_d and R'_d, respectively. We can write $\Delta t(T, R_d, \mathcal{W}^* \rightarrow \mathcal{W}'^*)$ and $\Delta t(T, R'_d, \mathcal{W}'^* \rightarrow \mathcal{W}^*)$ as follows according to (12):

$$\begin{aligned}
&\Delta t(T, R_d, \mathcal{W}^* \rightarrow \mathcal{W}'^*) \\
&\quad = R_d \cdot c_0 \cdot \sum_{E \in T} (w^*_E - w'^*_E) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E, E' \in T} F(E, E') \cdot \left(\frac{w^*_{E'}}{w^*_E} - \frac{w'^*_{E'}}{w'^*_E} \right) \\
&\qquad + \mathcal{K}_3 \cdot \sum_{E, E' \in T} F(E, E') \cdot \left(\frac{1}{w^*_E} - \frac{1}{w'^*_E} \right) \\
&\qquad + \mathcal{K}_4 \cdot \sum_{E \in T} G(E) \cdot \left(\frac{1}{w^*_E} - \frac{1}{w'^*_E} \right) \\
&\qquad + \mathcal{K}_5 \cdot \sum_{E \in T} H(E) \cdot \left(\frac{1}{w^*_E} - \frac{1}{w'^*_E} \right) \leq 0 \\
&\Delta t(T, R'_d, \mathcal{W}'^* \rightarrow \mathcal{W}^*) \\
&\quad = R'_d \cdot c_0 \cdot \sum_{E \in T} (w'^*_E - w^*_E) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E, E' \in T} F(E, E') \cdot \left(\frac{w'^*_{E'}}{w'^*_E} - \frac{w^*_{E'}}{w^*_E} \right) \\
&\qquad + \mathcal{K}_3 \cdot \sum_{E, E' \in T} F(E, E') \cdot \left(\frac{1}{w'^*_E} - \frac{1}{w^*_E} \right) \\
&\qquad + \mathcal{K}_4 \cdot \sum_{E \in T} G(E) \cdot \left(\frac{1}{w'^*_E} - \frac{1}{w^*_E} \right) \\
&\qquad + \mathcal{K}_5 \cdot \sum_{E \in T} H(E) \cdot \left(\frac{1}{w'^*_E} - \frac{1}{w^*_E} \right) \leq 0
\end{aligned}$$

Summing up the above two inequalities, we obtain

$$\begin{aligned}
&\Delta t(T, R_d, \mathcal{W}^* \rightarrow \mathcal{W}'^*) + \Delta t(T, R'_d, \mathcal{W}'^* \rightarrow \mathcal{W}^*) \\
&\quad = c_0 \cdot (R_d - R'_d) \cdot \sum_{E \in T} (w^*_E - w'^*_E) \leq 0
\end{aligned}$$

Since $R_d < R'_d$, we have $\sum_{E \in T} (w^*_E - w'^*_E) \geq 0$, i.e. area$(\mathcal{W}^*) \geq$ area$(\mathcal{W}'^*)$. □

Lemma 2: Consider an edge E in the routing tree. Let w^*_E and w'^*_E be the width of the edge in assignments $\mathcal{W}^*$ and $\mathcal{W}'^*$, respectively. If $w^*_E < w'^*_E$, then

$$\begin{aligned}
&\sum_{E' \in \mathrm{Ans}(E)} H(E') \cdot \left(\frac{1}{w^*_{E'}} - \frac{1}{w'^*_{E'}} \right) \\
&\qquad + \frac{H(E)}{w^*_E \cdot w'^*_E} \sum_{E' \in \mathrm{Des}(E)} (w'^*_{E'} - w^*_{E'}) \\
&\qquad > 0
\end{aligned} \tag{26}$$

Proof: Rewrite $t(T, R_d, \mathcal{W})$ in (12) with respect to E as follows:

$$\begin{aligned}
&t(T, R_d, \mathcal{W}) \\
&\quad = \mathcal{K}_0 + \left\{ R_d \cdot c_0 \cdot \sum_{E' \in T - \{E\}} w_{E'} + R_d \cdot c_0 \cdot w_E \right\} \\
&\qquad + \left\{ \mathcal{K}_2 \cdot \sum_{E', E'' \in T - \{E\}} F(E', E'') \cdot \frac{w_{E''}}{w_{E'}} \right. \\
&\qquad\quad + \mathcal{K}_2 \cdot \sum_{E' \in \mathrm{Ans}(E)} F(E', E) \cdot \frac{w_E}{w_{E'}} \\
&\qquad\quad \left. + \mathcal{K}_2 \cdot \sum_{E' \in \mathrm{Des}(E)} F(E, E') \cdot \frac{w_{E'}}{w_E} \right\} \\
&\qquad + \left\{ \mathcal{K}_3 \cdot \sum_{E' \in T - \{E\}, E'' \in T} F(E', E'') \cdot \frac{1}{w_{E'}} \right. \\
&\qquad\quad \left. + \mathcal{K}_3 \cdot \sum_{E' \in \mathrm{Des}(E)} F(E, E') \cdot \frac{1}{w_E} \right\} \\
&\qquad + \left\{ \mathcal{K}_4 \cdot \sum_{E' \in T - \{E\}} G(E') \cdot \frac{1}{w_{E'}} + \mathcal{K}_4 \cdot G(E) \cdot \frac{1}{w_E} \right\} \\
&\qquad + \left\{ \mathcal{K}_5 \cdot \sum_{E' \in T - \{E\}} H(E') \cdot \frac{1}{w_{E'}} + \mathcal{K}_5 \cdot H(E) \cdot \frac{1}{w_E} \right\}
\end{aligned}$$

Let $\mathcal{W}/E : w_E \rightarrow w'_E$ denote the wire .sizing solution obtained by replacing w_E with w'_E in $\mathcal{W}$. The difference $\Delta t(T, R_d, \mathcal{W}^*/E : w^*_E \rightarrow w'^*_E)$, i.e., $t(T, R_d, \mathcal{W}^*) - t(T, R_d, \mathcal{W}^*/E: w^*_E \rightarrow w'^*_E)$ is given by:

$$\begin{aligned}
&\Delta t(T, R_d, \mathcal{W}^*/E : w^*_E \rightarrow w'^*_E) \\
&\quad = R_d \cdot c_0 \cdot (w^*_E - w'^*_E) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E' \in \mathrm{Ans}(E)} F(E', E) \cdot \left(\frac{w^*_E - w'^*_E}{w^*_{E'}} \right) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E' \in \mathrm{Des}(E)} F(E, E') \cdot w^*_{E'} \cdot \left(\frac{1}{w^*_E} - \frac{1}{w'^*_E} \right) \\
&\qquad + \mathcal{K}_3 \cdot \sum_{E' \in \mathrm{Des}(E)} F(E, E') \cdot \left(\frac{1}{w^*_E} - \frac{1}{w'^*_E} \right) \\
&\qquad + \mathcal{K}_4 \cdot G(E) \cdot \left(\frac{1}{w^*_E} - \frac{1}{w'^*_E} \right) \\
&\qquad + \mathcal{K}_5 \cdot H(E) \cdot \left(\frac{1}{w^*_E} - \frac{1}{w'^*_E} \right) \\
&\quad \leq 0
\end{aligned}$$

Similarly, we can obtain $\Delta t(T, R'_d, \mathcal{W}'^*/E : w'^*_E \rightarrow w^*_E)$ and

deduce the following:

$$\begin{aligned}
&\Delta t(T, R_d, \mathcal{W}^*/E: w_E^* \to w_E'^*) \\
&\quad + \Delta t(T, R_d', \mathcal{W}'^*/E: w_E'^* \to w_E^*) \\
&\quad = (R_d - R_d') \cdot c_0 \cdot (w_E^* - w_E'^*) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E' \in \mathrm{Ans}(E)} F(E', E) \cdot (w_{E'}^* - w_{E'}'^*) \\
&\qquad \times \left(\frac{1}{w_{E'}^*} - \frac{1}{w_{E'}'^*} \right) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E' \in \mathrm{Des}(E)} F(E, E') \cdot (w_{E'}^* - w_{E'}'^*) \\
&\qquad \times \left(\frac{1}{w_E^*} - \frac{1}{w_E'^*} \right) \\
&\quad \le 0
\end{aligned}$$

Since $w_E^* - w_E'^* < 0$ and $F(E_1, E_2) = H(E_1)$ if $E_2 \in \mathrm{Des}(E_1)$,

$$\begin{aligned}
&(R_d - R_d') \cdot c_0 + \mathcal{K}_2 \cdot \sum_{E' \in \mathrm{Ans}(E)} H(E') \cdot \left(\frac{1}{w_{E'}^*} - \frac{1}{w_{E'}'^*} \right) \\
&\quad + \mathcal{K}_2 \cdot \sum_{E' \in \mathrm{Des}(E)} H(E) \cdot (w_{E'}'^* - w_{E'}^*) \cdot \left(\frac{1}{w_E^* \cdot w_E'^*} \right) \\
&\quad \ge 0
\end{aligned}$$

Since $R_d - R_d' < 0$ and $\mathcal{K}_2 > 0$, the result follows. □

As defined in [11] and [12], a *single-stem tree* is a tree with only one segment (called the *stem segment* of that tree) incident on its root.

Lemma 3: Consider a single stem tree T. Let E be the stem segment of tree T. If $R_d < R_d'$, then $w_E^* \ge w_E'^*$.

Proof: Suppose $w_E^* < w_E'^*$, we obtain the following from (26) in Lemma 2:

$$\frac{H(E)}{w_E^* \cdot w_E'^*} \cdot \sum_{E' \in \mathrm{Des}(E)} (w_{E'}'^* - w_{E'}^*) > 0$$

This implies that $\sum_{E' \in \mathrm{Des}(E)} w_{E'}'^* > \sum_{E' \in \mathrm{Des}(E)} w_{E'}^*$. Hence, area($\mathcal{W}'^*$) > area($\mathcal{W}^*$). However, this contradicts Lemma 1. □

Lemma 4: If $w_E^* < w_E'^*$ and $w_{E'}^* \ge w_{E'}'^*$ for all $E' \in \mathrm{Ans}(E)$, then

$$\sum_{E' \in \mathrm{Des}(E)} (w_{E'}'^* - w_{E'}^*) > 0 \tag{27}$$

Proof: Since $\sum_{E' \in \mathrm{Ans}(E)} H(E') \cdot (\frac{1}{w_{E'}^*} - \frac{1}{w_{E'}'^*}) \le 0$, it follows directly from (26) in Lemma 2 that the following is true:

$$H(E) \cdot \sum_{E' \in \mathrm{Des}(E)} (w_{E'}'^* - w_{E'}^*) > 0$$

Since $H(E) \ge 0$ (λ_i's are non-negative), it implies that $H(E) > 0$ and $\sum_{E' \in \mathrm{Des}(E)} (w_{E'}'^* - w_{E'}^*) > 0$. □

Lemma 5: Let E be an edge in the routing tree T. Let P_E denote the edges $E' \in \mathrm{Ans}(E)$ and T_E denote the single-stem subtree rooted at E (including E). Given R_d and the wire widths of edges in P_E, optimizing the wire width assignment of edges in $T - P_E$ implies that the following objective function is minimized:

$$\begin{aligned}
&t_{P_E}(T_E, R_d, \mathcal{W}) \\
&\quad = R_d \cdot c_0 \cdot \sum_{E' \in T_E} w_{E'} + \mathcal{K}_2 \cdot \sum_{E' \in P_E, E'' \in T_E} F(E', E'') \cdot \frac{w_{E''}}{w_{E'}} \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E', E'' \in T_E} F(E', E'') \cdot \frac{w_{E''}}{w_{E'}} \\
&\qquad + \mathcal{K}_3 \cdot \sum_{E', E'' \in T_E} F(E', E'') \cdot \frac{1}{w_{E'}} \\
&\qquad + \mathcal{K}_4 \cdot \sum_{E' \in T_E} G(E') \cdot \frac{1}{w_{E'}} + \mathcal{K}_5 \cdot \sum_{E' \in T_E} H(E') \cdot \frac{1}{w_{E'}}
\end{aligned} \tag{28}$$

Proof: Let $\overline{T_E \cup P_E}$ denote the edges not in $T_E \cup P_E$. We define:

$$\begin{aligned}
&t_{P_E}(P_E, R_d, \mathcal{W}) \\
&\quad = \sum_{E \in P_E} \mathcal{K}_1 \cdot w_E + \mathcal{K}_2 \cdot \sum_{E, E' \in P_E} F(E, E') \cdot \frac{w_{E'}}{w_E} \\
&\qquad + \mathcal{K}_3 \cdot \sum_{E \in P_E, E' \in T} F(E, E') \cdot \frac{1}{w_E} \\
&\qquad + \mathcal{K}_4 \cdot \sum_{E \in P_E} G(E) \cdot \frac{1}{w_E} \\
&\qquad + \mathcal{K}_5 \cdot \sum_{E \in P_E} H(E) \cdot \frac{1}{w_E} \\
&t_{P_E}(\overline{T_E \cup P_E}, R_d, \mathcal{W}) \\
&\quad = \mathcal{K}_1 \cdot \sum_{E' \in \overline{T_E \cup P_E}} w_{E'} \\
&\qquad + \mathcal{K}_2 \times \sum_{E' \in P_E, E'' \in \overline{T_E \cup P_E}} F(E, E') \cdot \frac{w_{E''}}{w_{E'}} \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E', E'' \in \overline{T_E \cup P_E}} F(E', E'') \cdot \frac{w_{E''}}{w_{E'}} \\
&\qquad + \mathcal{K}_3 \cdot \sum_{E', E'' \in \overline{T_E \cup P_E}} F(E', E'') \cdot \frac{1}{w_E'} \\
&\qquad + \mathcal{K}_4 \cdot \sum_{E' \in \overline{T_E \cup P_E}} G(E') \cdot \frac{1}{w_{E'}} \\
&\qquad + \mathcal{K}_5 \cdot \sum_{E' \in \overline{T_E \cup P_E}} H(E') \cdot \frac{1}{w_{E'}}
\end{aligned}$$

Hence, we can rewrite objectives function specified by (12) as

$$\begin{aligned}
t(T, R_d, \mathcal{W}) &= \mathcal{K}_0 + t_{P_E}(P_E, R_d, \mathcal{W}) \\
&\quad + t_{P_E}(\overline{T_E \cup P_E}, R_d, \mathcal{W}) + t_{P_E}(T_E, R_d, \mathcal{W})
\end{aligned}$$

Given the wire width assignment of P_E, $\mathcal{K}_0$ and $t_{P_E}(P_E, R_d, \mathcal{W})$ are both constants. Since the variables (widths of edges in T_E) in $t_{P_E}(T_E, R_d, \mathcal{W})$ do not appear in $t_{P_E}(\overline{T_E \cup P_E}, R_d, \mathcal{W})$, optimizing $t(T, R_d, \mathcal{W})$ implies

optimizing $t_{P_E}(T_E, R_d, \mathcal{W})$. (Note that the proof of Lemma 5 essentially proves the separability property in [9] and [13]). □

Lemma 6: If T is a single stem tree, then $R_d < R'_d$ implies $\mathcal{W}^*$ dominates $\mathcal{W}'^*$.

Proof: Sort all edges in T according to a topological order staring from the stem. Let E be the first edge in this order such that $w^*_E < w'^*_E$. Then, for all $E' \in \mathrm{Ans}(E), w^*_{E'} \geq w'^*_{E'}$. We use P_E to denote the edges $E' \in \mathrm{Ans}(E)$. By Lemma 3, P_E is not an empty set. Let T_E denote the subtree rooted at E (including E). Lemma 5 implies that the wire width assignment $\mathcal{W}^*$ of edges in T_E, denoted as $\mathcal{W}^*(T_E)$, is an optimal wire width assignment of T_E under the objective function (28), assuming that the wire widths of edges in P_E are fixed as specified by $\mathcal{W}^*$.

Let $\mathcal{W}^*/T_E : \mathcal{W}^* \rightarrow \mathcal{W}'^*$ denote the wire width assignment obtained by replacing the wire width assignment $\mathcal{W}^*(T_E)$ of T_E with $\mathcal{W}'^*(T_E)$, i.e., $(\mathcal{W}^* - \mathcal{W}^*(T_E)) \cup \mathcal{W}'^*(T_E)$. Similarly, let $\mathcal{W}'^*/T_E : \mathcal{W}'^* \rightarrow \mathcal{W}^*$ denote the wire width assignment obtained by replacing the wire width assignment $\mathcal{W}'^*(T_E)$ of T_E with $\mathcal{W}^*(T_E)$, i.e., $(\mathcal{W}'^* - \mathcal{W}'^*(T_E)) \cup \mathcal{W}^*(T_E)$. Comparing $t_{P_E}(T_E, R_d, \mathcal{W}^*)$ with $t_{P_E}(T_E, R_d, \mathcal{W}^*/T_E : \mathcal{W}^* \rightarrow \mathcal{W}'^*)$, we obtain the following inequality:

$$
\begin{aligned}
&t_{P_E}(T_E, R_d, \mathcal{W}^*) - t_{P_E}(T_E, R_d, \mathcal{W}^*/T_E : \mathcal{W}^* \rightarrow \mathcal{W}'^*) \\
&\quad = R_d \cdot c_0 \cdot \sum_{E' \in T_E} (w^*_{E'} - w'^*_{E'}) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E' \in P_E, E'' \in T_E} F(E', E'') \cdot \left(\frac{w^*_{E''}}{w^*_{E'}} - \frac{w'^*_{E''}}{w^*_{E'}} \right) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E', E'' \in T_E} F(E', E'') \cdot \left(\frac{w^*_{E''}}{w^*_{E'}} - \frac{w'^*_{E''}}{w'^*_{E'}} \right) \\
&\qquad + \mathcal{K}_3 \cdot \sum_{E', E'' \in T_E} F(E', E'') \cdot \left(\frac{1}{w^*_{E'}} - \frac{1}{w'^*_{E'}} \right) \\
&\qquad + \mathcal{K}_4 \cdot \sum_{E' \in T_E} G(E') \cdot \left(\frac{1}{w^*_{E'}} - \frac{1}{w'^*_{E'}} \right) \\
&\qquad + \mathcal{K}_5 \cdot \sum_{E' \in T_E} H(E') \cdot \left(\frac{1}{w^*_{E'}} - \frac{1}{w'^*_{E'}} \right) \\
&\quad \leq 0
\end{aligned}
$$

Similarly, comparing $t_{P_E}(T_E, R'_d, \mathcal{W}'^*)$ with $t_{P_E}(T_E, R'_d, \mathcal{W}'^*/T_E : \mathcal{W}'^* \rightarrow \mathcal{W}^*)$, we obtain the following inequality:

$$
\begin{aligned}
&t_{P_E}(T_E, R'_d, \mathcal{W}'^*) - t_{P_E}(T_E, R'_d, \mathcal{W}'^*/T_E : \mathcal{W}'^* \rightarrow \mathcal{W}^*) \\
&\quad = R'_d \cdot c_0 \cdot \sum_{E' \in T_E} (w'^*_{E'} - w^*_{E'}) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E' \in P_E, E'' \in T_E} F(E', E'') \cdot \left(\frac{w'^*_{E''}}{w'^*_{E'}} - \frac{w^*_{E''}}{w'^*_{E'}} \right) \\
&\qquad + \mathcal{K}_2 \cdot \sum_{E', E'' \in T_E} F(E', E'') \cdot \left(\frac{w'^*_{E''}}{w'^*_{E'}} - \frac{w^*_{E''}}{w^*_{E'}} \right) \\
&\qquad + \mathcal{K}_3 \cdot \sum_{E', E'' \in T_E} F(E', E'') \cdot \left(\frac{1}{w'^*_{E'}} - \frac{1}{w^*_{E'}} \right) \\
&\qquad + \mathcal{K}_4 \cdot \sum_{E' \in T_E} G(E') \cdot \left(\frac{1}{w'^*_{E'}} - \frac{1}{w^*_{E'}} \right) \\
&\qquad + \mathcal{K}_5 \cdot \sum_{E' \in T_E} H(E') \cdot \left(\frac{1}{w'^*_{E'}} - \frac{1}{w^*_{E'}} \right) \\
&\quad \leq 0
\end{aligned}
$$

Summing up the above two inequalities, we have:

$$
\begin{aligned}
&(R_d - R'_d) \cdot c_0 \cdot \sum_{E' \in T_E} (w^*_{E'} - w'^*_{E'}) \\
&\quad + \mathcal{K}_2 \cdot \sum_{E'' \in T_E} \left\{ (w^*_{E''} - w'^*_{E''}) \cdot \sum_{E' \in P_E} F(E', E'') \right. \\
&\qquad \left. \times \left(\frac{1}{w^*_{E'}} - \frac{1}{w'^*_{E'}} \right) \right\} \\
&\quad \leq 0
\end{aligned}
$$

Since $F(E', E'') = H(E')$ for $E' \in P_E$ and $E'' \in T_E$, we have:

$$
\begin{aligned}
&(R_d - R'_d) \cdot c_0 \cdot \sum_{E' \in T_E} (w^*_{E'} - w'^*_{E'}) \\
&\quad + \mathcal{K}_2 \cdot \sum_{E' \in P_E} H(E') \cdot \left(\frac{1}{w^*_{E'}} - \frac{1}{w'^*_{E'}} \right) \\
&\quad \times \sum_{E'' \in T_E} (w^*_{E''} - w'^*_{E''}) \leq 0
\end{aligned}
$$

From Lemma 4, $\sum_{E' \in T_E} (w'^*_{E'}) < 0$ [negation of (27)]. In addition, we know that $R_d - R'_d < 0$ and $\mathcal{K}_2 > 0$. Therefore, we can deduce that

$$
\sum_{E' \in P_E} H(E') \cdot \left(\frac{1}{w^*_{E'}} - \frac{1}{w'^*_{E'}} \right) > 0
$$

This contradicts the assumption that $w^*_{E'} \geq w'^*_{E'}$ for all $E' \in P_E$. □

According to the separability, we can decompose T into several single-stem trees and optimize each single-stem tree independently. Applying Lemma 6 to each single-stem tree, we have:

Theorem 5: (DS/WS Relation) For any tree T with one or more critical sinks, $R_d < R'_d$ implies $\mathcal{W}^*$ dominates $\mathcal{W}'^*$. □

This result reveals the relation between driver sizing and wiresizing, and it plays as important role in determining the lower and upper bounds of the optimal k-SDWS-D solutions in next subsection.

B. Lower and Upper Bound Computation for Optimal k-SDWS-D Solutions

We first present an algorithm called the *k-SDWS/D LU-Bound* algorithm for computing the upper and lower bounds of an optimal k-SDWS-D solution for a given number of stage k: Starting with an initial wire width assignment (say, all segments have the minimum width), we compute the

k-SDWS/D LU-Bound Algorithm

Function k-SDWS/D LU-Bound(T, R_{min}, k)

/* Given a tree T, the minimum width transistor resistance R_{min}, and the stage number k, return the upper and lower bounds of the optimal k-SDWS-D solution. */

$\mathcal{W}_l \leftarrow$ Min_Wire_Width;
while true
  Compute capacitive load: $C_{IL}(T, \mathcal{W}_l)$;
  Compute optimal driver stage ratio: $s_l \leftarrow (\frac{C_{IL}(T,\mathcal{W}_l)}{C_g})^{1/k}$;
  Compute the k-th driver resistance: $R_d \leftarrow \frac{R_{min}}{s_l^{(k-1)}}$;
  $\mathcal{W} \leftarrow$ Delay_Optimal_Wiresizing(R_d, T);
  if $\mathcal{W} > \mathcal{W}_l$ **then**
    $\mathcal{W}_l \leftarrow \mathcal{W}$
  else break;
end while
Output s_l as the stage ratio of the lower bound of driver sizing solution and $\mathcal{W}_l$ as the lower bound of wire sizing solution;
$\mathcal{W}_u \leftarrow$ Max_Wire_Width;
while true
  Compute capacitive load: $C_{IL}(T, \mathcal{W}_u)$;
  Compute optimal driver stage ratio: $s_u \leftarrow (\frac{C_{IL}(T,\mathcal{W}_u)}{C_g})^{1/k}$;
  Compute the k-th driver resistance: $R_d \leftarrow \frac{R_{min}}{s_u^{(k-1)}}$;
  $\mathcal{W} \leftarrow$ Delay_Optimal_Wiresizing(R_d, T);
  if $\mathcal{W} < \mathcal{W}_u$ **then**
    $\mathcal{W}_u \leftarrow \mathcal{W}$
  else break;
end while
Output s_u as the stage ratio of the upper bound of driver sizing solution and $\mathcal{W}_u$ as the upper bound of wire sizing solution;
end Function;

Fig. 4. The k-SDWS/D LU-Bound algorithm for computing lower and upper bounds of optimal SDWS solution for a given stage number k.

capacitive load of the routing tree. Based on Theorem 1, the optimal stage ratio, denoted by s, of the k-stage cascaded drivers can be computed and a driver sizing solution is obtained. Now, we perform delay optimal wiresizing [12] on the routing tree T based on the resistance of the k-th driver. If the wire width assignment changes, the capacitive load changes. A new driver stage ratio is computed to yield a new optimal driver sizing solution. Then, the optimal wiresizing solution will be computed again based on the new driver size of the k-th driver. The process is repeated until the wire width assignment does not change in consecutive iterations. The algorithm is described formally in Fig. 4.

Let $\mathcal{W}_0$ be the initial minimum width assignment and $\mathcal{W}_i$ denote the optimal wire width assignment obtained in iteration i. Similarly, let R_d^i denote the resistance of the k-th driver at the beginning of iteration i. During iteration i ($i \geq 1$), R_d^i is computed based on the capacitive load $C_{\mathrm{IL}}(T, \mathcal{W}_{i-1})$ and $\mathcal{W}_i$ is obtained based on R_d^i. It is obvious that in the first iteration, the wire widths in $\mathcal{W}_1$ is an upward revision of the wire width in $\mathcal{W}_0$ (since $\mathcal{W}_0$ is the minimum-width assignment). Hence, the capacitive load can only increase which implies an increase in the stage ratio and a decrease in the resistance of the k-th driver, i.e., $R_d^2 < R_d^1$. According to Theorem 5, $\mathcal{W}_2$ dominates $\mathcal{W}_1$. In general, we can show by mathematical induction that $R_d^{i+1} \leq R_d^i$ (since $\mathcal{W}_i$ dominates $\mathcal{W}_{i-1}$) and in turn $\mathcal{W}_{i+1}$ dominates $\mathcal{W}_i$ (according to Theorem 5). Hence, the algorithm will terminate (since there is an upper bound on the maximum wire width) and indeed computes the lower bound of driver and wire sizes of the optimal k-SDWS-D solution.

Similarly, we start with a maximum wire width assignment and compute the upper bound of driver and wire sizes of the optimal k-SDWS-D solution. Therefore, we have the following result:

Theorem 6: Given a chain of k drivers, the k-SDWS/D LU-Bound algorithm computes the lower and upper bounds of an optimal k-SDWS-D solution.

Proof: We give the proof for the lower bound computation only. The proof for upper bound can be derived similarly. Let $(\mathcal{D}^*, \mathcal{W}^*)$ be the optimal SDWS-D solution for a chain of k drivers and d_k^* be the optimal size of driver k. Let $(\mathcal{D}_i, \mathcal{W}_i)$ denote the driver and wiresizing solution computed in iteration i and d_k^i be the size of driver k in iteration i. We first show that $(\mathcal{D}_1, \mathcal{W}_1)$ is dominated by $(\mathcal{D}^*, \mathcal{W}^*)$.

Since $\mathcal{W}_0$ is dominated by $\mathcal{W}^*$, we have $C_{\mathrm{IL}}(T, \mathcal{W}_0) \leq C_{\mathrm{IL}}(T, \mathcal{W}^*)$. By Theorem 1,

$$d_k^1 = \left(\frac{C_{\mathrm{IL}}(T, \mathcal{W}_0)}{C_g}\right)^{(k-1)/k} \leq \left(\frac{C_{\mathrm{IL}}(T, \mathcal{W}^*)}{C_g}\right)^{(k-1)/k} = d_k^*$$

By Theorem 5, we conclude that $\mathcal{W}_1$ (which is computed based on R_d^1) is dominated by $\mathcal{W}^*$ since $R_d^1 = R_{\min}/d_k^1 \geq R_{\min}/d_k^* = R_d^*$ where R_d^* is the resistance of driver k in the optimal solution.

Assuming that $(\mathcal{D}_i, \mathcal{W}_i)$ is dominated by $(\mathcal{D}^*, \mathcal{W}^*)$, we can follow the same arguments and deduce that $(\mathcal{D}_{i+1}, \mathcal{W}_{i+1})$ is also dominated by $(\mathcal{D}^*, \mathcal{W}^*)$. Hence, the solution computed is a lower bound of the optimal solution for a given stage number k. □

Experimental results show that the algorithm terminates after three or four iterations in most cases. In addition, the upper and lower bounds meet for most instances, which implies that the optimal k-SDWS-D solution is obtained. Note that the upper and lower bounds include both the driver and wire sizes.

C. Optimal Algorithm for the SDWS-D Problem

For a given stage number k, we compute the optimal driver and wire sizing solution as follows: we compute the upper and lower bounds of the k-SDWS-D optimal solutions using the k-SDWS/D LU-Bound algorithm. In the case where the bounds do not meet, we can obtain a set of discrete driver sizes defined by the bounds computed by the k-SDWS/D LU-Bound algorithm for the k-th driver. For each driver size, we can compute the optimal wiresizing solution using the algorithm in [12] and use Theorem 1 to compute the sizes of driver $D_2 \cdots D_{k-1}$ using $d_k \cdot C_g$ as the capacitive load that $(k-1)$-stage drivers have to drive. The *k-SDWS/D Optimal* algorithm is given in Fig. 5.

Theorem 7: Given a chain of k drivers and p possible sizes for the last driver and a routing tree with n segments and r possible wire widths, the worst case time complexity to compute the optimal driver and wire sizes for the k-SDWS-D problem is $O(p \cdot n^r)$. □

Note that p is usually very small since the lower and upper bounds computed in Section III-B are very tight (in fact, they meet for almost all test cases). The factor $O(n^r)$ is the worst case complexity of optimal wiresizing algorithm which in fact

k-SDWS/D Optimal Algorithm

```
Function k-SDWS/D Optimal(T, R_min, k)
    /* Given a tree T, the minimum width transistor resistance R_min, and a
    stage number k, return the optimal stage ratio s, and the optimal wire width
    assignment W. */
    Compute the upper and lower bounds:
        {s_u, s_l, W_u, W_l} ← k-SDWS/D LU-Bound(T, R_min, k);
    if W_u > W_l (upper and lower bounds do not meet) then
        Best_Delay ← ∞;
        Discretize possible driver sizes of the k-th driver in the interval
            [(s_l)^{k-1}, (s_u)^{k-1}];
        for each driver size d_p in [(s_l)^{k-1}, (s_u)^{k-1}] do
            d_k ← d_p;
            Compute k-th driver resistance: R_d ← R_min / d_p
            W ← Delay_Optimal_Wiresizing(R_d, T);
            Compute driver sizes for D_{1..k-1} according to Theorem 1;
            Current_Delay ← t(T, k, D, W);
            if Current_Delay < Best_Delay then
                Solution ← {D, W};
                Best_Delay ← Current_Delay;
        end for
    else /* W_u = W_l and s_u = s_l */
        Construct Solution from {s_u, s_l, W_u, W_l};
        Best_Delay ← t(T, k, D, W);
    end if
    return Solution;
end Function;
```

Fig. 5. The k-SDWS/D Optimal algorithm for computing the optimal SDWS solution for a given stage number k.

SDWS/D Optimal Algorithm

```
Function SDWS/D Optimal(T, R_min)
    /* Given a tree T, and the minimum width transistor resistance R_min, return
    the optimal number of stages k, the optimal stage ratio s, and the optimal
    wire width assignment W. */
    Compute k^D_MAX;
    Best_Delay ← ∞;
    for k ← 1 to k^D_MAX do
        Compute optimal driver and wire sizing solution for k:
            {D, W} ← k-SDWS/D Optimal(T, R_min, k);
        Current_Delay ← t(T, k, D, W);
        if Current_Delay < Best_Delay then
            Best_Solution ← {k, D, W};
    end for
    return Best_Solution;
end Function;
```

Fig. 6. The SDWS/D Optimal algorithm.

run in $O(n^3 \cdot r)$ time based on effective lower and upper bound computation using the dominance property (Details of the wiresizing algorithm can be found in [12]). Therefore, the k-SDWS/D Optimal algorithm runs in $O(n^3 \cdot r)$ time in practice.

To compute the optimal stage number k_D^* for the SDWS-D problem, our *SDWS/D Optimal* algorithm performs a linear search for $1 \le k \le k_{\mathrm{MAX}}^D = \left\lceil \frac{\ln C_{\mathrm{IL}}(T, \mathcal{W}_{\mathrm{MAX}})/C_g}{\ln s^*} \right\rceil$ where s^* is the optimal stage ratio to be defined later in Lemma 8. The algorithm is given in Fig. 6. Optimality of the SDWS/D Optimal algorithm is justified by the following results:

Lemma 7: Given a routing tree T with a fixed wire width assignment $\mathcal{W}$, the performance measure $t(T, k, \mathcal{D}, \mathcal{W})$ in (15) is a concave function with respect to k for $k > 0$ (assuming k is continuous).

Proof: We can rewrite the performance measure $t(T, k, \mathcal{D}, \mathcal{W})$ in (15) as

$$t(T,k,\mathcal{D},\mathcal{W}) = k \cdot R_{\min} \cdot C_d + R_{\min} \cdot C_g \cdot \sum_{i=1}^{k-1} \frac{d_{i+1}}{d_i} + R_{\min} \cdot \frac{C_{\mathrm{IL}}(T,\mathcal{W})}{d_k} + t(T,\mathcal{W}) \qquad (29)$$

where

$$t(T,\mathcal{W}) = \sum_{N_i \in \mathrm{sink}(T)} \lambda_i \cdot \sum_{E \in P(N_+, N_i)} \frac{r_0 \cdot c_0}{2} + \mathcal{K}_2 \cdot \sum_{E,E' \in T} F(E,E') \cdot \frac{w_{E'}}{w_E} + \mathcal{K}_3 \cdot \sum_{E,E' \in T} F(E,E') \cdot \frac{1}{w_E} + \mathcal{K}_4 \cdot \sum_{E \in T} G(E) \cdot \frac{1}{w_E} + \mathcal{K}_5 \cdot \sum_{E \in T} H(E) \cdot \frac{1}{w_E}$$

Note that $t(T, \mathcal{W})$ is a constant for a fixed wire width assignment.

Given a fixed wire width assignment $\mathcal{W}$, we can compute $C_{\mathrm{IL}}(T, \mathcal{W})$. Using Theorem 1, $t(T, k, \mathcal{D}, \mathcal{W})$ is minimized when the stage ratio is $s = (\frac{C_{\mathrm{IL}}(T,\mathcal{W})}{C_g})^{1/k}$ for a given k. Let $M = (\frac{C_{\mathrm{IL}}(T,\mathcal{W})}{C_g})$, we can rewrite the minimized performance measure as

$$t(T,k,\mathcal{D},\mathcal{W}) = k \cdot R_{\min} \cdot C_d + k \cdot R_{\min} \cdot C_g \cdot M^{1/k} + t(T,\mathcal{W})$$

Taking the second derivative of $t(T, k, \mathcal{D}, \mathcal{W})$ with respect to k, we obtain

$$\frac{d^2 t(T,k,\mathcal{D},\mathcal{W})}{dk^2} = R_{\min} \cdot C_g \cdot \frac{1}{k^3} \cdot M^{1/k} \cdot \ln M > 0$$

Note that $M > 1$ since $C_{\mathrm{IL}}(T, \mathcal{W})$ is assumed to be larger than C_g. Hence, $t(T, k, \mathcal{D}, \mathcal{W})$ is a concave function. □

Since $t(T, \mathcal{W})$ is a constant in (29) for a fixed wire width assignment, the following result holds:

Lemma 8: [21] Given a routing tree with a fixed wire width assignment $\mathcal{W}$, the optimal stage ratio s^* is given by the expression

$$s^* = e^{1+a/s^*}$$

where $a = C_d/C_g$ and e is the base of the natural logarithm. The corresponding optimal stage number $k_{\mathcal{W}}^*$, which may not be an integer, is

$$k_{\mathcal{W}}^* = \frac{\ln C_{\mathrm{IL}}(T,\mathcal{W})/C_g}{\ln s^*}$$

Note that s^* depends only on the process parameters and not the loading capacitance whereas $k_{\mathcal{W}}^*$ depends on both the process parameters and wire width assignment. Moreover, in our SDWS/D Optimal algorithm, $k_{\mathrm{MAX}}^D = \lceil k_{\mathcal{W}_{\mathrm{MAX}}}^* \rceil$.

Theorem 8: (Optimality of SDWS/D Optimal Algorithm): Given a routing tree T, the optimal stage number k_D^* is less than or equal to k_{MAX}^D.

Proof: Suppose not, i.e., $k_D^* > k_{\mathrm{MAX}}^D$. Let $\mathcal{W}^*$ denote the wiresizing solution of the optimal SDWS-D problem. Since $\mathcal{W}^*$ is dominated by $\mathcal{W}_{\mathrm{MAX}}, C_{\mathrm{IL}}(T, \mathcal{W}_{\mathrm{MAX}}) \geq C_{\mathrm{IL}}(T, \mathcal{W}^*)$. Therefore,

$$k_D^* > k_{\mathrm{MAX}}^D = \lceil k_{\mathcal{W}_{\mathrm{MAX}}}^* \rceil \geq \lceil k_{\mathcal{W}^*}^* \rceil \geq k_{\mathcal{W}^*}^*$$

From Lemmas 7 and 8, $t(T, k, \mathcal{D}, \mathcal{W})$ is a concave function and it is strictly monotonically increasing for $k > k_{\mathcal{W}}^*$. This contradicts the assumption that k_D^* is the optimal stage number since $k_D^* > \lceil k_{\mathcal{W}^*}^* \rceil \geq k_{\mathcal{W}^*}^*$. □

IV. Simultaneous Driver and Wire Sizing for Both Delay and Power Optimization (SDWS-DP Problem)

In the SDWS-DP problem formulated in Section II-B, the drivers and wires are sized to minimize the following objective function (22):

$$\mathrm{obj}(T, k, \mathcal{D}, \mathcal{W}) = \alpha \cdot \mathrm{Power}(T, k, \mathcal{D}, \mathcal{W}) + \gamma \cdot t(T, k, \mathcal{D}, \mathcal{W})$$

We can adjust parameters α and γ to achieve smooth trade-off between performance and power. By choosing the parameters carefully, we can compute a driver and wire sizing solution to meet a delay constraint while minimizing the power dissipation.

In this section, we shall first study the properties of the optimal SDWS-DP solutions and then present an efficient algorithm for computing the optimal SDWS-DP solution.

A. Properties of Optimal SDWS Solutions for Both Delay and Power Optimization

The lemmas and theorems in this subsection are used to prove the correctness and optimality of the bound computation and optimal algorithm to the SDWS-DP problem. If the reader is only interested to learn the SDWS/DP algorithms, one may proceed directly to Section IV-B.

First, we define the monotone property and the dominance relation[4] for the driver sizing solutions.

Definition 4: Given a chain of k cascaded drivers D_i's, a driver sizing solution $\mathcal{D}$ is a monotone assignment if $d_{i+1} > d_i$ for all $i = 1, 2, \ldots, k-1$.

Definition 5: Given two driver sizing solutions $\mathcal{D} = \{d_1, d_2, \ldots, d_k\}$ and $\mathcal{D}' = \{d_1', d_2', \ldots, d_k'\}$, $\mathcal{D}$ dominates $\mathcal{D}'$ if $d_i \geq d_i'$ for $1 \leq i \leq k$.

The following theorem defines a set of equations which will be used for optimal driver sizing solution computation for a fixed loading capacitance $C_{\mathrm{IL}}(T, \mathcal{W})$ in Section IV-B.

Theorem 9: Given a fixed loading capacitance $C_{\mathrm{IL}}(T, \mathcal{W})$, the optimal driver size assignment $\mathcal{D}$ for a chain of k drivers satisfies the following equations when $\gamma > 0$:

$$\begin{gathered}\alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) + \gamma \cdot \mathcal{J}_2 \cdot \left(\frac{1}{d_{i-1}} - \frac{d_{i+1}}{d_i^2}\right) = 0 \\ \text{for all } i = 2 \ldots k-1 \\ \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) + \gamma \cdot \mathcal{J}_2 \cdot \left(\frac{1}{d_{k-1}} - \frac{C_{\mathrm{IL}}(T, \mathcal{W})}{d_k^2 \cdot C_g}\right) = 0\end{gathered} \tag{30}$$

[4] It has been shown in [9] that the dominance property for driver sizing solutions, which can be defined in a similar manner as the dominance property for wiresizing solution in Theorem 4, holds.

Proof: The optimal driver sizing solution minimizes the objective function specified by (23):

$$\mathrm{obj}_{T,k,\mathcal{W}}(\mathcal{D}) = \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2}^{k} d_i + \gamma \cdot \mathcal{J}_2 \cdot \left(\frac{C_{\mathrm{IL}}(T, \mathcal{W})}{d_k \cdot C_g} + \sum_{i=1}^{k-1} \frac{d_{i+1}}{d_i}\right)$$

Differentiating $\mathrm{obj}_{T,k,\mathcal{W}}(\mathcal{D})$ with respect to d_i for $2 \leq i \leq k$, we obtain

$$\begin{aligned}\frac{\partial \mathrm{obj}_{T,k,\mathcal{W}}(\mathcal{D})}{\partial d_i} &= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) + \gamma \cdot \mathcal{J}_2 \cdot \left(\frac{1}{d_{i-1}} - \frac{d_{i+1}}{d_i^2}\right) \\ &\quad \text{for all } i = 2 \cdots k-1 \\ \frac{\partial \mathrm{obj}_{T,k,\mathcal{W}}(\mathcal{D})}{\partial d_k} &= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) + \gamma \cdot \mathcal{J}_2 \cdot \left(\frac{1}{d_{k-1}} - \frac{C_{\mathrm{IL}}(T, \mathcal{W})}{d_k^2 \cdot C_g}\right)\end{aligned}$$

By setting $\frac{\partial \mathrm{obj}_{T,k,\mathcal{W}}(\mathcal{D})}{\partial d_i} = 0$ for $2 \leq i \leq k$, we obtain the system of equations given in (30). □

Note that there are many solutions to the set of nonlinear equations specified by (30) and they may assign drivers with sizes smaller than the minimum driver size. The following result allows us to consider only monotone drive sizing solution.

Theorem 10: (Monotone Property) For any given capacitive load $C_{\mathrm{IL}}(T, \mathcal{W})$, any optimal driver sizing solution to the SDWS-DP problem under the combined delay and power optimization objective function specified by (23) is monotone.

Proof: Let $\mathcal{D}$ be the nonmonotone optimal driver size assignment. Suppose j is the smallest index such that $d_j \leq d_{j-1}$. Let $\mathcal{D} - \{d_j\}$ be the assignment obtained after discarding driver D_j in the nonmonotone optimal driver chain. For convenience, we use d_{k+1} to denote $\frac{C_{\mathrm{IL}}(T, \mathcal{W})}{C_g}$. According to objective function defined in (23), we have:

$$\begin{aligned}\mathrm{obj}_{T,k,\mathcal{W}}(\mathcal{D}) &= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2}^{k} d_i + \gamma \cdot \mathcal{J}_2 \cdot \sum_{i=1}^{k} \frac{d_{i+1}}{d_i} \\ \mathrm{obj}_{T,k-1,\mathcal{W}}(\mathcal{D} - \{d_i\}) &= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2, i \neq j}^{k} d_i + \gamma \cdot \mathcal{J}_2 \\ &\quad \times \left(\sum_{i=1}^{j-2} \frac{d_{i+1}}{d_i} + \frac{d_{j+1}}{d_{j-1}} + \sum_{i=j+1}^{k} \frac{d_{i+1}}{d_i}\right)\end{aligned}$$

Therefore,

$$\Delta = \text{obj}_{T,k,\mathcal{W}}(\mathcal{D}) - \text{obj}_{T,k-1,\mathcal{W}}(\mathcal{D} - \{d_i\})$$
$$= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot d_j + \gamma \cdot \mathcal{J}_2 \cdot \left(\frac{d_j}{d_{j-1}} + \frac{d_{j+1}}{d_j} - \frac{d_{j+1}}{d_{j-1}} \right)$$

Since $\frac{1}{d_j} - \frac{1}{d_{j-1}} \geq 0$, we have $\Delta > 0$, which contradicts the assumption that $\mathcal{D}$ is optimal. □

We shall show the following result which allows us to apply a similar approach as that in Section III to solve the SDWS-DP problem:

Theorem 11: (WS/DS Relation) Given two wire width assignments $\mathcal{W}$ and $\mathcal{W}'$ for a routing tree T driven by a chain of k drivers. Let $\mathcal{D}$ and $\mathcal{D}'$ be the optimal driver sizing solutions for $\mathcal{W}$ and $\mathcal{W}'$, respectively. If $C_{\text{IL}}(T, \mathcal{W}) \geq C_{\text{IL}}(T, \mathcal{W}')$, then $\mathcal{D}$ dominates $\mathcal{D}'$.

Proof: Let $\mathcal{D} = \{d_1, d_2, \ldots, d_k\}$ and $\mathcal{D}' = \{d'_1, d'_2, \ldots, d'_k\}$. First we show that $d_k \geq d'_k$.

$$\text{obj}_{T,k,\mathcal{W}}(\mathcal{D}) - \text{obj}_{T,k,\mathcal{W}}(\mathcal{D}')$$
$$= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2}^{k} (d_i - d'_i) + \gamma \cdot \mathcal{J}_2$$
$$\times \left(\frac{C_{\text{IL}}(T, \mathcal{W})}{C_g} \cdot \left(\frac{1}{d_k} - \frac{1}{d'_k} \right) + \sum_{i=1}^{k-1} \left(\frac{d_{i+1}}{d_i} - \frac{d'_{i+1}}{d'_i} \right) \right)$$
$$\leq 0$$

$$\text{obj}_{T,k,\mathcal{W}'}(\mathcal{D}') - \text{obj}_{T,k,\mathcal{W}'}(\mathcal{D})$$
$$= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2}^{k} (d'_i - d_i) + \gamma \cdot \mathcal{J}_2$$
$$\times \left(\frac{C_{\text{IL}}(T, \mathcal{W}')}{C_g} \cdot \left(\frac{1}{d'_k} - \frac{1}{d_k} \right) + \sum_{i=1}^{k-1} \left(\frac{d'_{i+1}}{d'_i} - \frac{d_{i+1}}{d_i} \right) \right)$$
$$\leq 0$$

Summing up the above two inequalities, we obtain:

$$(C_{\text{IL}}(T, \mathcal{W}) - C_{\text{IL}}(T, \mathcal{W}')) \cdot \left(\frac{1}{d_k} - \frac{1}{d'_k} \right) \leq 0$$

Since $C_{\text{IL}}(T, \mathcal{W}) \geq C_{\text{IL}}(T, \mathcal{W}')$, we conclude that $d_k \geq d'_k$.

Given $1 < j \leq k$, suppose $d_i \geq d'_i$ for $j \leq i \leq k$. We want to show that $d_{j-1} \geq d'_{j-1}$ and hence prove the dominance relation by induction. Let $\mathcal{D}/d_{1\cdots j-1} \rightarrow d'_{1\cdots j-1}$ denote the driver size assignment by replacing the sizes of drivers $D_{1\cdots j-1}$ in $\mathcal{D}$ by the corresponding sizes in $\mathcal{D}'$. Similarly, we define $\mathcal{D}'/d'_{1\cdots j-1} \rightarrow d_{1\cdots j-1}$ to be the driver size assignment by replacing the sizes of drivers $D_{1\cdots j-1}$ in $\mathcal{D}'$ by the corresponding sizes in $\mathcal{D}$.

$$\text{obj}_{T,k,\mathcal{W}}(\mathcal{D}) - \text{obj}_{T,k,\mathcal{W}}(\mathcal{D}/d_{1\cdots j-1} \rightarrow d'_{1\cdots j-1})$$
$$= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2}^{j-1} (d_i - d'_i) + \gamma \cdot \mathcal{J}_2$$
$$\times \left\{ \sum_{i=1}^{j-2} \left(\frac{d_{i+1}}{d_i} - \frac{d'_{i+1}}{d'_i} \right) + \left(\frac{d_j}{d_{j-1}} - \frac{d_j}{d'_{j-1}} \right) \right\} \leq 0$$

$$\text{obj}_{T,k,\mathcal{W}'}(\mathcal{D}') - \text{obj}_{T,k,\mathcal{W}'}(\mathcal{D}'/d'_{1\cdots j-1} \rightarrow d_{1\cdots j-1})$$
$$= \alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot \sum_{i=2}^{j-1} (d'_i - d_i) + \gamma \cdot \mathcal{J}_2$$
$$\times \left\{ \sum_{i=1}^{j-2} \left(\frac{d'_{i+1}}{d'_i} - \frac{d_{i+1}}{d_i} \right) + \left(\frac{d'_j}{d'_{j-1}} - \frac{d'_j}{d_{j-1}} \right) \right\} \leq 0$$

Summing up the above two inequalities, we obtain:

$$(d_j - d'_j) \cdot \left(\frac{1}{d_{j-1}} - \frac{1}{d'_{i-1}} \right) \leq 0$$

since $d_j - d'_j \geq 0$, we have $d_{i-1} \geq d'_{i-1}$. □

Furthermore, one can see that Formulation (24) implies that Theorems 2 through 5 in Sections III-A-2 and III-A-3 still apply to the optimal wiresizing solution under the combined delay and power optimization objective as defined in (24). Therefore, the same optimal wiresizing solution algorithm in [12] can be used to optimize $\text{obj}_{T,k,\mathcal{D}}(\mathcal{W})$ in (24).

B. Bound Computation and Optimal Algorithm for SDWS-DP Problem

For a given stage number k, we taken the same approach as in solving the k-SDWS-D problem to compute the optimal solutions to the k-SDWS-DP problem: we first compute the upper and lower bounds of the optimal k-SDWS-DP solution. To compute the lower bound, we start with an initial wire width assignment in which all segments have minimum width wire. Based on the capacitive load of the routing tree, we compute a driver sizing solution using (30). Then, the optimal wiresizing solution based on the current driver sizing solution is computed using the algorithm in [12]. The process of alternative driver sizing and wiresizing is repeated until the wire sizing solutions do not change in consecutive iterations. The upper bound is computed similarly by starting with maximum wire width assignment. The algorithm outlined above is referred to as the *k-SDWS/DP LU-Bound* algorithm. Note that the driver sizing solution given by k-SDWS/DP LU-Bound algorithm is computed by solving (30) while the driver sizing solution given by k-SDWS/D LU-Bound algorithm is computed based on Theorem 1. We have the following result (similar to the SDWS-D problem):

Theorem 12: Given a chain of k drivers, the k-SDWS/DP LU-Bound algorithm computes the lower and upper bounds of an optimal k-SDWS-DP solution.

Proof: We use the same notations introduced in the proof of Theorem 6. Again, we prove the lower bound case only. We first show that $(\mathcal{D}_1, \mathcal{W}_1)$ is dominated by $(\mathcal{D}^*, \mathcal{W}^*)$.

Since $\mathcal{W}_0$ is dominated by $\mathcal{W}^*$, we have $C_{\text{IL}}(T, \mathcal{W}_0) \leq C_{\text{IL}}(T, \mathcal{W}^*)$. By Theorem 11 (instead of Theorem 1), $\mathcal{D}_1$ is dominated by $\mathcal{D}^*$, which implies that $d^1_k \leq d^*_k$. According to Theorem 5, $\mathcal{W}_1$ (computed based on R^1_d) is dominated by $\mathcal{W}^*$ since $R^1_d = R_{\min}/d^1_k \geq R_{\min}/d^*_k = R^*_d$, where R^*_d is the resistance of driver k in the optimal solution.

Assuming that $(\mathcal{D}_i, \mathcal{W}_i)$ is dominated by $(\mathcal{D}^*, \mathcal{W}^*)$, we can show similarly that $(\mathcal{D}_{i+1}, \mathcal{W}_{i+1})$ is also dominated by $(\mathcal{D}^*, \mathcal{W}^*)$. Hence, according to mathematical induction, the

solution computed is a lower bound of the optimal solution for a given drive stage number k. □

Given the lower and upper bounds $[d_k^l, d_k^u]$ of the k-th driver, the *k-SDWS/DP Optimal* algorithm tries every possible driver size in $[d_k^l, d_k^u]$ for the k-th driver size. Again, (30) (instead of Theroem 1) is used to compute the sizes of driver $D_2 \cdots D_{k-1}$ using $d_k \cdot C_g$ as the capacitive load that the $(k-1)$-stage drivers have to drive. For each possible d_k, the optimal wiresizing solution can still be computed using the algorithm in [12]. As in the SDWS-D problem, we have $d_k^l = d_k^u$ for most cases and a very small interval $[d_k^l, d_k^u]$ when $d_k^l \neq d_k^u$. We establish the following result which is similar to Theorem 7:

Theorem 13: Given a chain of k drivers and p possible sizes for the last driver and a routing tree with n segments and r possible wire widths, let $T(k)$ be the time taken to compute the sizes of drivers $D_2 \cdots D_{k-1}$ given the sizes of D_k by solving the systems of equations in (30). Then, the worst case time complexity to compute the optimal driver and wire sizes is $O(p \cdot (n^r + T(k)))$. □

Again, the factor $O(n^r)$ is the worst case time complexity for computing an optimal wiresizing solution. In practice, when dominance property is used to compute the lower and upper bounds of the optimal wiresizing solution, this term is reduced to $O(n^3 \cdot r)$ for almost all designs.

Due to the similarity between (15) and (24), optimal wiresizing solution under the combined objective of performance and power dissipation can be computed using the algorithm in [12]. However, computing an optimal driver sizing solution for a given capacitive load $C_{\text{IL}}(T, \mathcal{W})$ is more difficult. In the SDWS-D algorithms, the driver sizing solution is obtained according to Theorem 1 by computing a fixed stage ratio. In the SDWS-DP algorithms, we have to solve the set of equations specified in (30). In our implementation, these equations are solved using MAPLE, a mathematical software for symbolic computation developed by University of Waterloo.

To obtain the optimal stage number k_{DP}^*, the *SDWS/DP Optimal* algorithm performs a search for $1 \leq k < k_{\text{MAX}}^{\text{DP}}$ where $k_{\text{MAX}}^{\text{DP}}$ is the smallest stage number such that capacitive load $C_{\text{IL}}(T, \mathcal{W}_{\text{MAX}})$ for a routing tree T with maximum wire width assignment $\mathcal{W}_{\text{MAX}}$ does not have a monotone optimal solution for a chain of $k_{\text{MAX}}^{\text{DP}}$ drivers. The following results prove the correctness of the SDWS-DP Optimal algorithm.

Lemma 9: For a given capacitive load C_{IL} driven by a chain of $k \geq 2$ drivers, if there exists a monotone optimal k-SDWS-DP driver sizing solution, then there exists a monotone optimal k-SDWS-DP driver sizing solution for any capacitive load $C'_{\text{IL}} > C_{\text{IL}}$. Equivalently, if there exists no monotone optimal k-stage driver sizing solution to the k-SDWS-DP problem for capacitive load C_{IL}, then there exists no monotone optimal k-SDWS-DP driver sizing solution for any capacitive load $C'_{\text{IL}} > C_{\text{IL}}$.

Proof: It is trivially true for $k = 2$ (Theorem 11). Suppose this is true for $k = l$. Now, consider $k = l + 1$. Let d_{l+1} and d'_{l+1} be the optimal driver size for driver D_{l+1} driving C_{IL} and C'_{IL} respectively. According to Theorem 11, $d_{l+1} \leq d'_{l+1}$. By hypothesis, there exists a monotone optimal driver solution for a chain of l drivers driving a capacitive load of $d'_{l+1} \cdot C_g$. We only have to show that $d'_{l+1} > d'_l$. By (30),

$$d_l'^2 = (d'_{l-1} \cdot d'_{l+1}) \Big/ \left(\frac{\alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot d'_{l-1}}{\gamma \cdot \mathcal{J}_2} + 1 \right) < d'_{l-1} \cdot d'_{l+1}$$

Since $d'_l > d'_{l-1}$, we can conclude that $d'_{l+1} > d'_l$. Hence, there exists a monotone optimal driver sizing solution for a chain of k drivers. The other result follows directly. □

Lemma 10: For a given capacitive load $C_{\text{IL}}(T, \mathcal{W})$, if there exists no monotone optimal driver sizing solution for a chain of k drivers, then there exists no monotone optimal driver sizing solution for a chain of $k + 1$ drivers.

Proof: Suppose there exists a monotone optimal driver sizing solution for a chain of $k+1$ drivers. Let $\mathcal{D}$ and $\mathcal{D}'$ denote the optimal driver sizing solution for the chain of k drivers and $k+1$ drivers, respectively. Note that $\mathcal{D}$ is a nonmonotone solution whereas $\mathcal{D}'$ is a monotone solution. First we show that $d'_{k+1} < \frac{C_{\text{IL}}(T,\mathcal{W})}{C_g}$. According to (30),

$$d_{k+1}'^2 = \left(d'_k \cdot \frac{C_{\text{IL}}(T, \mathcal{W})}{C_g} \right) \Big/ \left(\frac{\alpha \cdot (\mathcal{L}_1 + \mathcal{L}_2) \cdot d'_k}{\gamma \cdot \mathcal{J}_2} + 1 \right) < d'_k \cdot \frac{C_{\text{IL}}(T, \mathcal{W})}{C_g}$$

Since $d'_k < d'_{k+1}$, we conclude that $d'_{k+1} < \frac{C_{\text{IL}}(T,\mathcal{W})}{C_g}$. Since d'_i's for $1 \leq i \leq k+1$ are monotone, there exists a monotone optimal solution for a chain of k drivers driving any capacitive load larger than $d'_{k+1} \cdot C_g$ (according to Lemma 9). However, this contradicts the assumption that there exists no monotone optimal solution for a chain of k drivers driving capacitive load $C_{\text{IL}}(T, \mathcal{W})$. □

Theorem 14: (Optimality of SDWS/DP Optimal Algorithm): Let $k_{\text{MAX}}^{\text{DP}}$ be the smallest stage number such that capacitive load $C_{\text{IL}}(T, \mathcal{W}_{\text{MAX}})$ for a routing tree T with maximum wire width assignment $\mathcal{W}_{\text{MAX}}$ does not have a monotone optimal solution for a chain of $k_{\text{MAX}}^{\text{DP}}$ drivers. There exists no monotone optimal driver sizing solution for any $k \geq k_{\text{MAX}}^{\text{DP}}$ cascaded drivers given any wire width assignments.

Proof: By Lemma 10, there exists no monotone optimal k-stage driver sizing solution for capacitive load $C_{\text{IL}}(T, \mathcal{W}_{\text{MAX}})$ for any $k \geq K_{\text{MAX}}^{\text{DP}}$. Since $C_{\text{IL}}(T, \mathcal{W}_{\text{MAX}}) \geq C_{\text{IL}}(T, \mathcal{W})$ for any wire width assignment $\mathcal{W}$, we can conclude from Lemma 9 that there exists no monotone optimal k-stage driver sizing solution for $k \geq k_{\text{MAX}}^{\text{DP}}$ given any wiresizing solution (including the optimal wiresizing solution). □

To compute $k_{\text{MAX}}^{\text{DP}}$, we perform a linear search starting with stage number $k = 2$. By solving the system of equations given in (30) at each stage, we can verify if a monotone solution exists for the current stage number.[5] We define $k_{\text{MAX}}^{\text{DP}}$ to be the first k such that the monotone property does not hold. By Lemma 10, there exists no monotone k-SDWS-DP optimal driver sizing solution for any $k > k_{\text{MAX}}^{\text{DP}}$ cascaded drivers driving the capacitive load $C_{\text{IL}}(T, \mathcal{W}_{\text{MAX}})$. Since all possible wire width assignments are dominated by $\mathcal{W}_{\text{MAX}}$,

[5] By using the *k-monotone* property defined in [9], we can speed up the verification process.

SDWS/DP Optimal Algorithm

```
Function SDWS/DP Optimal(T, R_min, α, γ)
    /* Given a tree T, the minimum width transistor resistance R_min, and the
    parameters α and γ, return the optimal number of stages k, the optimal
    driver sizing solution D, and the optimal wire width assignment W. */
    if α = 0 then
        return SDWS/D Optimal(T, R_min); /* Delay Optimization */
    if γ = 0 then
        return k = 1, d_1 = 1, W = W_MIN; /* Minimum Power Dissipation */
    Compute k_MAX^DP;
    Best_Combined_Objective ← ∞;
    for k ← 1 to k_MAX^DP − 1 do
        Compute optimal driver and wire sizing solution for k:
            {D, W} ← k-SDWS/DP Optimal(T, R_min, k, α, γ);
        Current_Combined_Objective ← obj(T, k, D, W);
        if Current_Combined_Objective < Best_Combined_Objective then
            Best_Solution ← {k, D, W};
    end for
    return Best_Solution;
end Function;
```

Fig. 7. The SDWS/DP Optimal algorithm.

TABLE I
TECHNOLOGY PARAMETERS FOR CAZM 0.5 μm CMOS DRIVERS [18]

Min Driver Resistance (Ω)	13598
Min Gate Capacitance (fF)	2.6802
Min Diffusion Capacitance (fF)	1.0403

TABLE II
TECHNOLOGY PARAMETERS BASED ON (a) MCM10 TECHNOLOGY [5] AND (b) CAZM 0.5 μm CMOS MODE [18]

Parameters	MCM	IC
Min Loading Capacitance (fF)	1000	2.6802
Wire Resistance ($\Omega/\Box$)	0.02	0.044
Wire Capacitance (area) ($aF/\mu m^2$)	3.46	41.3
Fringing Capacitance (2 sides) ($aF/\mu m$)	50.4	150

there exists no monotone optimal driver sizing solution for any $k > k_{\mathrm{MAX}}^{\mathrm{DP}}$ cascaded drivers given any wire width assignments. Therefore, it is sufficient for the SDWS/DP Optimal algorithm to compute, for each $1 \leq k < k_{\mathrm{MAX}}^{\mathrm{DP}}$, the optimal wire and driver sizing solution (using the k-SDWS/DP Optimal algorithm). The best among the $k_{\mathrm{MAX}}^{\mathrm{DP}} - 1$ solutions is the optimal SDWS-DP solution. The SDWS/DP Optimal algorithm is given in Fig. 7.

V. EXPERIMENTAL RESULTS

We have implemented the optimal SDWS/D and SDWS/DP algorithms in ANSI C for the Sun SPARC station environment. The algorithm is tested on a number of MCM and IC design examples consisting of a chain of cascaded drivers driving 1) a single sink net through a long interconnect and 2) a multiple-sink net through a tree-structure interconnect. HSPICE is used to simulate the circuit for accurate timing and power simulation to verify both our models and algorithms. The technology parameters used by HSPICE are summarized in Table I and II.

The parameters for the driver are based on the CAZM 0.5 μm CMOS process model [18]. The minimum driver resistance is obtained for a minimum size transistor (width = 1.0 μm, length = 0.5 μm) through HSPICE simulation of the drain-to-source current available when the drain-to-source voltage for a n-device is 0.95 $\times V_{\mathrm{dd}}$ (V_{dd} = 5.0 V). The minimum gate and diffusion capacitance is the gate and diffusion capacitance of a minimum size n-transistor. We assume in the experiments that drivers in all design examples (MCM and IC) are cascaded on-chip CMOS drivers. The minimum driver resistance, gate and diffusion capacitance are used to guide the design of drivers.

The minimum loading capacitance on a MCM substrate, which is the pad capacitance, is 1000 fF whereas the minimum loading capacitance on an IC chip is the minimum gate capacitance. The interconnect parameters are obtained from MCM10 model [5] and CAZM 0.5 μm CMOS process model [18].

A. Performance of SDWS/D Algorithm

In our experiments, we compare our SDWS/D solutions with the solutions obtained using a) driver sizing with constant stage ratio $s = e$ and minimum-wire-width (CDSMIN), b) optimal driver sizing based on Theorem 1 and minimum-wire-width (ODSMIN), and c) independent driver and wire sizing algorithm (DWSA), i.e., it performs driver sizing with constant stage ratio $s = e$ and optimal wiresizing [8]. The set of wire width allowed is $\{W_1, 2W_1, 3W_1, 4W_1\}$, where W_1 is the minimum width (10 μm in MCM10 technology and 0.95 μm in IC technology). Hence, every wire segment in CDSMIN and ODSMIN has width W_1.

For the experiment on MCM design examples, we assume that in each test net, the cascaded drivers are driving a) a sink through a 5 cm long interconnect, b) a 4-sink net randomly placed on 10 cm $\times$ 10 cm substrate, and c) a 8-sink net randomly placed on 10 cm $\times$ 10 cm substrate. In each case, the interconnect is divided into wire segments, each of length 100 μm and modeled by a π-type RC circuit, in order to model the distributed nature of interconnect. For each test circuit in b) and c), we performed two sets of experiments. In the first set of experiments, we let the farthest sink from the source be the critical sink. In the second set of experiments, we randomly generate two critical sinks for each 4-sink net and four critical sinks for each 8-sink net. The first driver in the chain is in turn driven by an ideal voltage source and the input signal is a square wave with rise and fall time of 1 ns and a period of 40 ns (25 MHz).

Table III summarizes the signal delays by the four algorithms. In both sets of experiments, our SDWS/D algorithm consistently outperforms the other three methods. Compared with the traditional methods of constant-ratio driver sizing or optimal driver sizing with uniform wire width (CDSMIN and ODSMIN), our SDWS/D solutions achieved an improvement of up to 49% in both cases. Compared with the recently reported DWSA method [8], our SDWS/D algorithm can reduce the delay by up to 11%.

We perform the same set of experiments on the following IC design examples: a) a sink through a 1 cm long interconnect, b) a 4-sink net randomly placed on 1 cm $\times$ 1 cm IC chip, and c) a 8-sink net randomly placed on 1 cm $\times$ 1 cm IC chip. In the IC

TABLE III
AVERAGE SIGNAL DELAY (ns) FOR (a) 1-CRITICAL-SINK NET AND (b) MULTICRITICAL-SINKS NET UNDER THE MCM10 TECHNOLOGY

Net	Single Critical Sink				Multiple Critical Sinks			
	CDSMIN	ODSMIN	DWSA	SDWS/D	CDSMIN	ODSMIN	DWSA	SDWS/D
1-sink	1.31	1.17	1.15	1.02	same as single critical sink			
4-sink	4.36	4.38	2.41	2.26	4.35	4.33	2.45	2.31
8-sink	5.54	5.53	2.82	2.80	4.05	3.63	2.24	2.14

TABLE IV
AVERAGE SIGNAL DELAY (ns) FOR (a) 1-CRITICAL-SINK NET AND (b) MULTICRITICAL-SINK NETS UNDER THE CAZM 0.5 μm CMOS TECHNOLOGY

Net	Single Critical Sink				Multiple Critical Sinks			
	CDSMIN	ODSMIN	DWSA	SDWS/D	CDSMIN	ODSMIN	DWSA	SDWS/D
1-sink	1.54	1.48	1.05	0.95	same as single critical sink			
4-sink	2.21	2.13	1.35	1.19	2.14	2.05	1.28	1.17
8-sink	2.53	2.47	1.44	1.30	1.91	1.83	1.22	1.09

TABLE V
POWER DISSIPATION (mW) AND DELAY (ns) FOR DIFFERENT α'S UNDER THE MCM10 TECHNOLOGY

α	$k=2$		$k=3$		$k=4$		$k=5$		$k=6$	
	Power	Delay	Power	Delay	Power	Delay	Power	Delay	Power	Delay
0	31.42	9.075	81.08	3.787	128.5	2.785	174.9	2.488	211.0	2.411
0.1	28.89	9.076	69.56	3.863	100.3	2.808	109.2	2.525	113.2	2.417
0.2	28.07	9.077	51.66	4.156	74.31	3.010	95.02	2.648	97.09	2.545
0.3	27.79	9.121	47.90	4.448	65.18	3.212	76.79	2.772	78.14	2.761
0.4	27.02	9.352	43.43	4.750	58.11	3.444	65.87	3.148	66.92	3.186
0.5	26.00	9.756	39.49	5.298	50.71	3.870	59.70	3.278	59.32	3.449
0.6	25.03	10.085	36.30	5.751	45.25	4.439	51.19	3.852	51.39	3.938
0.7	23.29	10.929	33.44	6.407	40.09	5.100	45.39	4.256	45.82	4.498
0.8	20.22	13.178	29.08	7.724	36.70	5.478	39.81	5.190	40.05	5.218
0.9	16.64	16.696	25.63	9.103	33.63	6.244	33.63	6.505	34.74	6.601

design, we segment the interconnect into grid edges of length 10 μm and the sink capacitance is 10 times the minimum loading capacitance. The results are summarized in Table IV. Again, our SDWS/D algorithm consistently outperforms the other three methods. Compared with the CDSMIN, ODSMIN and DWSA methods, our SDWS/D solutions achieved an improvement of up to 49, 47, and 12%, respectively.

B. Performance of SDWS/DP Algorithm

We have studied the trade-off between power dissipation and delay using our SDWS/DP algorithm. For the same set of MCM 4-sink nets (with two critical sinks) in the previous experiment, we compute for each stage number $k = 2$ to 6, a driver and wire sizing solution optimizing the $\text{obj}_\alpha(T, k, \mathcal{D}, \mathcal{W})$ for different α in (25). Table V summarizes the results for a set of α parameters. For each α, the table lists the power dissipation (mW) and delay (ns) of the driver and wire sizing solution for each stage number $k = 2$ to 6. As one can see, by adjusting the value of α in objective function (25), one can achieve smooth trade-off between power dissipation and performance.

In Fig. 8(a), the trade-off between delay and power dissipation for different k's is depicted by a set of curves. We can obtain a *contour* named *D/P-curve* which is dominated by the set of curves from above. Clearly, the D/P-curve represents best power and delay trade-off achieved by the optimal solutions of the SDWS/DP solutions. In Fig. 8(b), we compare the solutions obtained by CDSMIN, ODSMIN and DWSA with the D/P-curve by the SDWS/DP algorithm. We can observe that the D/P-curve is dominated by the curves representing the solution sets by CDSMIN, ODSMIN and DWSA. One can always select a SDWS solution from the D/P-curve to meet a delay target with minimum power dissipation or to meet a power budget with best performance.

TABLE VI
COMPARISONS OF POWER REQUIREMENT (mW) OF VARIOUS SOLUTIONS MEETING THE DELAY SPECIFICATION (ns) FOR MCM DESIGN

Delay Constraint (*ns*)	CDSMIN			ODSMIN			DWSA			SDWS-DP		
	Power	Delay	k	Power	Delay	k	Power	Delay	k	Power	Delay	k
10	29	7.67	5	29	10.00	2	30	7.31	5	26	9.10	3
5	44	4.99	7	83	4.63	4	54	3.86	6	43	4.74	3
4	-	-	-	-	-	-	54	3.86	6	50	3.87	4
3	-	-	-	-	-	-	86	2.69	7	77	2.77	5
2.5	-	-	-	-	-	-	139	2.45	8	113	2.41	6

TABLE VII
COMPARISONS OF POWER REQUIREMENT (mW) OF VARIOUS SOLUTIONS MEETING THE DELAY SPECIFICATION (ns) FOR IC DESIGN

Delay Constraint (*ns*)	CDSMIN			ODSMIN			DWSA			SDWS-DP		
	Power	Delay	k	Power	Delay	k	Power	Delay	k	Power	Delay	k
5	7.7	3.69	4	8.3	4.11	2	8.0	3.66	4	6.6	4.47	2
4	7.7	3.69	4	15.0	2.48	3	8.0	3.66	4	7.0	4.00	3
3	11.9	2.31	5	15.0	2.48	3	13.8	1.87	5	8.8	2.76	3
2	-	-	-	32.3	2.05	6	13.8	1.87	5	11.8	2.00	4
1.5	-	-	-	-	-	-	23.2	1.37	6	16.8	1.47	4

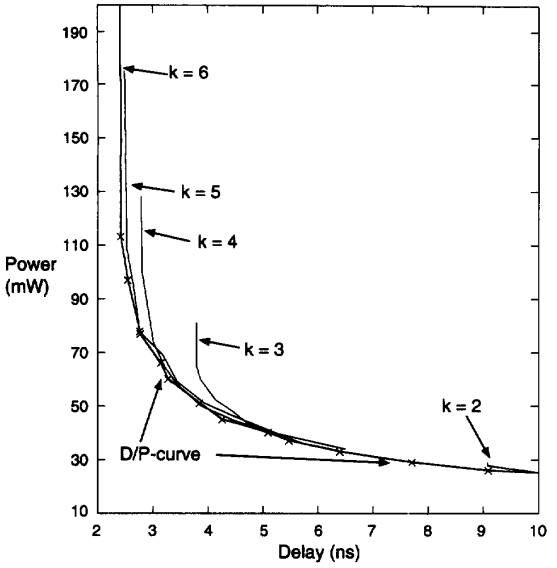
In Table VI, we compare the power requirement of the test circuit under different design methods (CDSMIN, ODSMIN, DWSA, and SDWS-DP) in order for the net to meet the delay specification. The table list a set of target delays that the net is expected to achieve and the most power economical solution by each design method under the performance requirement. An empty entry implies that the solution cannot meet the delay specification. We can observe that our SDWS solutions require least power while meeting the delay specification. Our SDWS solutions achieved an reduction in power dissipation by up to 10, 48, and 19% reduction, respectively, when compared with the CDSMIN, ODSMIN and DWSA methods. In addition, our SDWS solutions always require fewer stages of drivers (smaller k) which results in simpler layout design in practice.

We also evaluate the performance of our SDWS/DP algorithms on 4-sink nets (with two critical sinks) in IC routing used in the previous subsection. Table VII compares the power requirement of the test circuits under different design methods (CDSMIN, ODSMIN, DWSA and SDWS-DP) in order for the net to meet the delay specification. Our SDWS/DP algorithms can achieve up to 26, 63, and 36% reduction in power as compared to the CDSMIN, ODSMIN and DWSA methods, respectively.

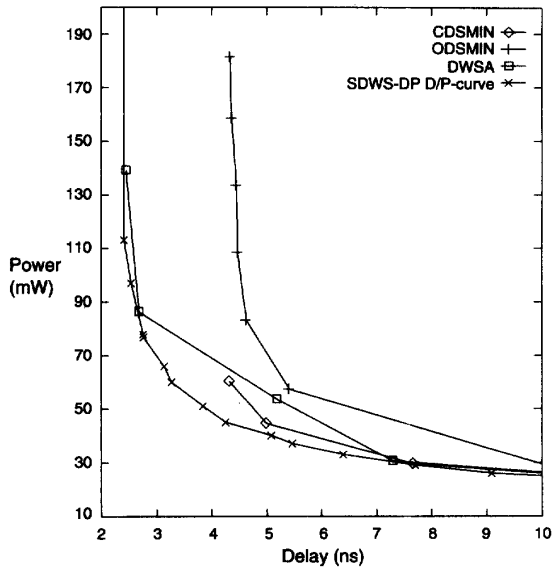
We can observe that by selecting α carefully, we can minimize power dissipation while meeting the delay constraint. This suggests an approach to solve the following optimization problem which was raised in [8]: Given a routing tree T and a performance specification B, determine the number of stages k, the set of driver sizes $\mathcal{D}$, and a wiresizing solution $\mathcal{W}$ on T, such that the performance measure $t(T, k, \mathcal{D}, \mathcal{W})$ is bounded by B, i.e., $t(T, k, \mathcal{D}, \mathcal{W}) \leq B$ and the power dissipation $\text{Power}(T, k, \mathcal{D}, \mathcal{W})$ is minimized. We can solve the above problem by performing a binary search on α in (25) where $0 \leq \alpha \leq 1$.

VI. CONCLUSION AND FUTURE WORK

To the best of our knowledge, this is the first paper which presents in-depth study of the simultaneous driver and wire sizing problem and its effects on performance and power op-



(a)



(b)

Fig. 8. (a) D/P-curve containing the optimal SDWS/DP solutions. (b) Comparisons between CDSMIN, ODSMIN, DWSA and SDWS/DP solutions under the MCM10 technology.

timization. The results in this paper have shown convincingly that simultaneous driver and wire sizing can lead to significant reduction in the interconnect delay and power dissipation. In high performance systems where the delays due to long on-chip interconnect and chip-to-chip interconnects are critical, our optimal solutions allow smaller driver size and/or fewer number of cascaded drivers required to drive long interconnect with proper wiresizing. As a result, our SDWS solutions provide a low power interconnect design for high performance circuits.

One of the limitations of our formulation is the simplicity of the *RC* model for drivers. It has been shown that the *RC* delay model may have error up to 25% [19]. We would like to use a more accurate model, such as the slope model or the PR-slope model proposed in [19]. In addition, we would like to extend the short-circuit power formulation (20) to model the rise/fall time of the input signal to individual driver along the chain and investigate if the monotone and dominance properties, the WS/DS relation, and the optimality of the SDWS/DP Optimal algorithm still hold under such model.

We would also like to generalize the driver sizing and wiresizing problem such that long interconnect lines can be segmented and interleaved with optimal size repeaters. Also, as mentioned in Section I, interconnect topology optimization is another effective approach to reduce signal delay. A long term goal is to include wiresizing, driver sizing and also interconnect topology in the formulation of the power and performance optimization problem.

ACKNOWLEDGMENT

The authors would like to thank Dr. D. Zhou for providing the technology parameters for CAZM 0.5 μm CMOS model, and anonymous reviewers for their helpful comments.

REFERENCES

[1] C. J. Alpert, T. C. Hu, J. H. Huang and A. B. Kahng, "A direct combination of the Prim and Dijkstra constructions for improved performance-driven routing," in *Proc. IEEE Int. Symp. on Circuits and Syst.*, May 1993, pp. 1869–1872.

[2] H. B. Bakoglu, *Circuits, Interconnections, and Packaging for VLSI*. Reading, MA: Addison-Wesley, 1990.

[3] K. D. Boese, A. B. Kahng, and G. Robins, "High-performance routing trees with identified critical sinks," in *Proc. ACM/IEEE Design Automation Conf.*, 1993, pp. 182–187.

[4] K. D. Boese, A. B. Kahng, B. A. McCoy, and G. Robins, "Rectilinear Steiner trees with minimum Elmore delay," in *Proc. ACM/IEEE Design Automation Conf.*, 1994, pp. 381–386.

[5] J. B. Burr, J. R. Burnham, and A. M. Peterson, "System-wide energy optimization in the MCM environment," in *Proc. IEEE MCM Conf.*, 1991, pp. 66–83.

[6] J. P. Cohoon and L. J. Randall, "Critical net routing," in *Proc. IEEE Int. Conf. on Computer Design*, 1991, pp. 174–177.

[7] J. Cong, A. B. Kahng, G. Robins, M. Sarrafzadeh, and C. K. Wong, "Provably good performance-driven global routing," *IEEE Trans. CAD*, vol. 11, no. 6, pp. 739–752, June 1992.

[8] J. Cong, C.-K. Koh, and K. S. Leung, "Wiresizing with driver sizing for performance and power optimization," in *Proc. 1994 Int. Workshop on Lower Power Design*, 1994, pp. 81–86.

[9] J. Cong and C.-K. Koh, "Simultaneous driver and wire sizing for performance and power optimization," UCLA Computer Science Dept. *Tech. Rep. CSD-940020*, Los Angeles, CA 90024, May 1994.

[10] ——, "Simultaneous driver and wire sizing for performance and power optimization," in *Proc. IEEE Int. Conf. on Computer-Aided Design*, 1994, pp. 206–212.

[11] J. Cong, K. S. Leung, and D. Zhou, "Performance-driven interconnect design based on distributed RC delay model," in *Proc. ACM/IEEE Design Automation Conf.*, 1993, pp. 606–611.

[12] J. Cong and K. S. Leung, "Optimal wiresizing under the distributed Elmore delay model," in *Proc. IEEE Int. Conf. on Computer-Aided Design*, 1993, pp. 634–639.

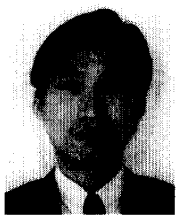
[13] ——, "Optimal wiresizing under Elmore delay model," to appear in *IEEE Trans. Computer-Aided Design*, Feb. 1995 (also available as UCLA Computer Science Dept. *Tech. Rep. CSD-930012*, Los Angeles, CA 90024, April 1993).

[14] W. C. Elmore, "The transient response of damped linear network with particular regard to wideband amplifier," *J. Appl. Phys.*, vol. 19, pp. 55–63, 1948.
[15] T. D. Hodes, B. A. McCoy, and G. Robins, "Dynamically-wiresized Elmore-based routing constructions," in *Proc. IEEE Int. Symp. Circuits and Syst.*, June 1994, pp. 463–466.
[16] X. Hong, T. Xue, E. S. Kuh, C. K. Cheng, and J. Huang, "Performance-driven Steiner tree algorithms for global routing," in *Proc. ACM/IEEE Design Automation Conf.*, 1993, pp. 177–181.
[17] H. C. Lin and L. W. Linholm, "An optimized output stage for MOS integrated circuits," *IEEE J. Solid-State Circ.*, vol. SC-10, no. 2, pp. 106–109, 1975.
[18] *MCNC Designers' Manual*, MCNC.
[19] J. K. Ousterhout, "Switch-level delay models for digital MOS VLSI," in *Proc. ACM/IEEE Design Automation Conf.*, 1984, pp. 542–548.
[20] S. S. Sapatnekar, "RC interconnect optimization under the Elmore delay model," in *Proc. ACM/IEEE Design Automation Conf.*, 1994, pp. 387–391.
[21] N. H. E. Weste and K. Eshraghian, *Principles of CMOS VLSI Design: a Systems Perspective*, 2nd ed. Reading, MA: Addison-Wesley, 1993, pp. 229–231.

Jason Cong (M'90) received the B.S. degree in computer science from the Peking University in 1985, and the M.S. and Ph.D. degrees in computer science from the University of Illinois at Urbana-Champaign in 1987 and 1990, respectively.

Currently, he is an associate professor with the Computer Science Department of University of California, Los Angeles. From 1986 to 1990, he was a research assistant with the Computer Science Department, University of Illinois. He worked at the Xerox Palo Alto Research Center in the summer of 1987 and at the National Semiconductor Corporation in the summer of 1988. He has served on the program committees of a number of VLSI CAD Conferences, and chaired the Fourth ACM/SIGDA Physical Design Workshop. His research interests include computer-aided design of VLSI circuits, rapid system prototyping, and design and analysis of combinatorial and geometric algorithms.

Dr. Cong received the Best Graduate Award from the Peking University in 1985 and the Ross J. Martin Award for Excellence in Research from the University of Illinois at Urbana-Champaign in 1989. He also received the National Science Foundation Research Initiation Award in 1991 and the National Science Foundation Young Investigator Award in 1993. He is also the recipient of the Northrop Corporation Outstanding Junior Faculty Research Award from the UCLA in 1993.



Cheng-Kok Koh received the B.S. degree in computer science with first class honors from the National University of Singapore in 1992.

Currently, he is a research assistant with the Department of Computer Science of University of California, Los Angeles, where he is pursuing the Ph.D. degree. His research interests include computer-aided design of VLSI circuits, design and analysis of data structures and algorithms, and hypermedia systems.

Mr. Koh is a member of the Association for Computing Machinery.

Methods to Improve Digital MOS Macromodel Accuracy

Jeong-Taek Kong and David Overhauser, *Member, IEEE*

Abstract— **This paper presents accurate series-transistor reduction techniques which extend the applicability of linear and nonlinear macromodels to more complex structures through accurately modeling the channel length modulation effect, effective transconductance, input terminal position dependence, parasitic capacitances, such as gate coupling capacitances, and the body effect. Adequate solutions to address these sources of delay errors, which may total 100% or more, have not been previously provided. The significant improvement in simulation accuracy using these proposed techniques is shown. The timing macromodel used to implement these techniques is up to several hundred times faster than SPICE2 and up to several times faster than existing nonlinear macromodels. The accuracy of this macromodel over a wide range of operating conditions is demonstrated. The macromodel and reduction techniques can be used to minimize VLSI simulation time, provide fast feedback in circuit optimization, and generate accurate data for higher-level macromodels. The proposed reduction techniques apply to linear and nonlinear macromodels.**

I. Introduction

TO IMPROVE the performance of electrical simulators, such as SPICE2 [1], for VLSI circuits, many different techniques have been applied. Most are based on relaxation techniques. These relaxation-based simulators are usually one order of magnitude faster than SPICE2, but this is still too slow for VLSI circuit verification (they are effective for circuits with at most up to several tens of thousands of transistors). Thus, switch-level timing and fast timing simulators attempt to bridge the gap between electrical simulators and logic simulators. Logic simulators are fast but generally inaccurate [2]. Switch-level timing and fast timing simulation apply a wide variety of techniques, such as macromodeling, to span this gap. The need for macromodeling has grown with the advent of VLSI. Even though simulation and macromodeling techniques are maturing, there is still a strong need for more powerful macromodeling techniques. As process and device technologies continuously evolve, the number of devices in VLSI circuits increases and the relative accuracy of existing macromodels decreases. In addition, macromodeling operation assumptions applied to previous technologies, such as lumping coupling capacitances and using a single equivalent transconductance for series-connected transistors, do not apply to advanced technologies.

There are five basic types of timing macromodeling techniques in use: 1) table lookup (or tabular) delay methods; 2) empirical delay equations; 3) *RC* delay and Asymptotic Waveform Evaluation (AWE) methods [3]; 4) inverter analysis techniques; and 5) nonlinear macromodeling techniques. The above techniques, used in a wide variety of simulators (i.e., logic, switch-level timing, fast timing simulators), are studied in detail in [2]. Both the table lookup delay methods and empirical delay equations have a common drawback which is the long time required for precharacterization of primitive subcircuits with a variety of input slopes and output loadings. *RC* delay and AWE methods oversimplify a transistor as a linear resistor, which produces inaccurate results. On the other hand, inverter analysis techniques analytically solve the differential equations at a subcircuit output node (i.e., closed-form output waveform or delay equations are obtained). In general, inverter analysis techniques use a scaling factor to map logic gates to a standard inverter. The main advantage of these techniques is that they do not require presimulation. Despite inaccuracies of the analytic inverter analysis with a step input and simple quadratic transistor equations, it has been widely used in real designs [4] as well as in most digital design texts [5]–[7] since its introduction in 1964 [8]. Its popularity is due to the simplicity of the analysis. An analytical solution for the CMOS inverter output response to a ramp input with quadratic transistor equations was first presented in 1987 [9]. Similar inverter analysis techniques have been reported in [10] and [11]. Recently, an improved model [12] has been derived in which simple quadratic transistor equations are used and the solution is obtained using approximation. However, analytical derivation of the output response is very difficult and is impossible in most cases when the input waveform is a complicated function of time or the transistor model equations are not simple.

To overcome some of the above limitations, nonlinear macromodeling techniques have been developed [13]–[17]. However, the nonlinear macromodeling techniques have a fundamental limitation in that they are still based on simple quadratic transistor equations, such as the SPICE level 1 model. It is well known that simple quadratic equations are not capable of simulating transistors accurately [18]. Table-

Manuscript received October 29, 1993; revised December 13, 1994. This work was supported in part by a fellowship from Samsung Electronics Company and the Semiconductor Research Corporation under Contract 93-DJ-305. This paper was recommanded by Associate Editor R. A. Saleh.

J.-T. Kong was with the Department of Electrical Engineering, Duke University, Durham, NC 27706 USA. He is now with CAE Group, Semiconductor Business, Samsung Electronics Company, Korea.

D. Overhauser is with the Department of Electrical Engineering, Duke University, Durham, NC 27706 USA.

IEEE Log Number 9411261.

based macromodels based on piecewise-linear interpolation of I_{ds} with respect to v_{ds} and v_{gs} [17] and piecewise-quadratic approximation of the transistor current equation [19] address the limitation of simple quadratic equations. Because the time step of the table-based macromodel in [17] is limited by a circuit simulation in the mixed-mode simulation environment, the accuracy of the macromodel can be maintained; however, the speedup may not be significant. If used in certain operating conditions with a larger time step, this approximation may not be accurate.

There have been several approaches to macromodeling based on more complicated transistor equations. The output response of an ED-MOS inverter is analytically derived in [20] where the driver transistor current is modeled by a bulk-charge model (i.e., similar to a simplified SPICE level 2 model) and the current of the load transistor is modeled using a quadratic equation (i.e., the SPICE level 1 model). In [20] the input waveform is represented as a staircase with a given number of fixed time steps Δt. This modeling technique applies an iteration scheme at each Δt. The analytic delay method with the nth-power law MOSFET model was proposed in [18]. A CMOS inverter with a rising ramp input is approximated as an n-channel transistor with an input waveform which remains zero until the input reaches the logic threshold voltage and matches the ramp input after that point. A similar approach, the hyperbolic tangent (tanh) law MOSFET model, was proposed in [21]. The tanh law MOSFET model was applied to the analysis of a CMOS inverter where both the p- and n-channel transistors were included in the differential equation at the output node, but the output response was solved numerically in [21].

Although existing macromodeling techniques can produce accurate results for generic MOS primitives [13]–[17] the issue of mapping series-connected transistors to the MOS primitive has not been adequately addressed. Existing macromodels use a single equivalent transconductance, β_{eq}, for series-connected transistors. Unfortunately, this simple approximation generates large errors in certain situations. These errors become clear when multiple input gates are examined. Fig. 1 shows the input position dependence of the response of a 4-input CMOS NAND gate. The solid curve corresponds to the topmost transistor switching. Macromodels which apply a single equivalent β produce results to the right of the solid curve, as shown by the dotted curve in Fig. 1. This approximation overestimates the delay by as much as 100% (i.e., too much pessimistic estimation). The output waveforms for 4 different inputs were obtained using SPICE2 and the conventional model curve was obtained using a nonlinear macromodel which applies a single equivalent β [15], [19].

In addition to finding β_{eq} for a generic MOS primitive, an equivalent channel length modulation factor, λ_{eq}, is also determined [15], [16]. λ_{eq} is determined using a simple linear scaling in [15] and [16]. An improved λ_{eq} is presented in this paper.

Another source of delay error is the inaccurate consideration of parasitic capacitances, such as coupling capacitances. Although [14]–[17] include coupling capacitances in the generic MOS primitive, it has not been implemented in [14]–[16]

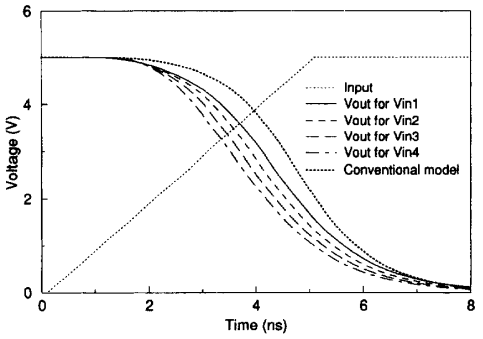


Fig. 1. Input position dependence of a 4-input CMOS NAND gate.

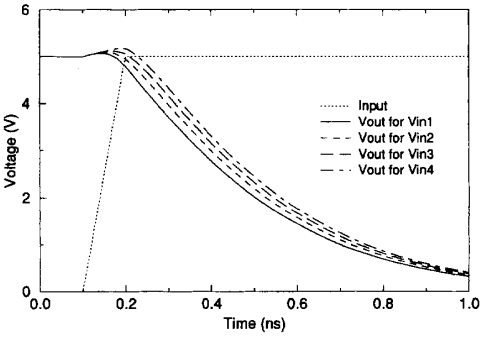


Fig. 2. Input position dependence on parasitic capacitances.

TABLE I
THE POTENTIAL DELAY ERROR INTRODUCED BY SECONDARY EFFECTS

secondary parameter	potential delay error	when large errors occur
β_{eq}	~ 100%	slow input and bottom transistor switching
coupling capacitances	~ 20%	fast input and small output capacitance
λ_{eq}	~ 10%	large λ

and the performance of the macromodel in [17] has not been shown. In [12], the input-output coupling capacitances in an inverter are modeled as constant capacitances. However, the impact of coupling capacitances is also dependent on input position. This position dependence, determined using SPICE2, is shown in Fig. 2. Traditional models underestimate the delay due to ignoring the effect of coupling capacitances. The solid curve corresponds to the topmost transistor switching. This dependence is also addressed in this paper.

The potential delay error introduced by failing to address these secondary effects is shown in Table I. The magnitude of error is a function of operating conditions, and thus may not always be evident. Existing β_{eq} estimates generate extra delay, and ignoring coupling capacitances underestimates delay. This paper introduces techniques to address these sources of error and shows how the accuracy of macromodeling can be significantly improved. These techniques apply to both linear and nonlinear macromodels, but are demonstrated in a nonlinear macromodel outlined in Section II. Section III addresses the mapping of series-connected transistors to the MOS primitive

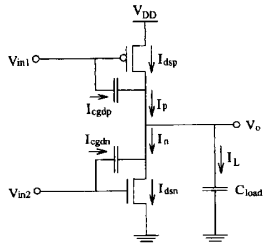


Fig. 3. Some components of a MOS macromodel.

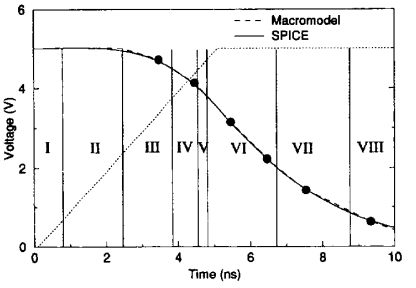


Fig. 4. Simulation results of a CMOS inverter. Six model evaluations are performed here. However, the tail portion does not need model evaluation in typical cases.

and Section IV illustrates the accuracy improvement using the new techniques.

II. A Nonlinear Macromodel

Macromodeling for timing simulation consists of two important parts: analysis of the primitive macromodel and reduction of complex gates to that macromodel. Although this work primarily addresses the second part, this section introduces the nonlinear macromodel on which the techniques of Section III are implemented. Further detail is found in [22].

Although nonlinear macromodels are generally applied to a generic MOS primitive [13]–[17], the most common form of the primitive is a configuration which resembles an inverter. Therefore, this form is referred to as an inverter macromodel although the techniques presented apply to the generic MOS primitive as well. The output response of the inverter is derived in several regions of operation. The total output load capacitance, C_L, is the sum of the effective lumped capacitance of parasitics and the output load capacitance, C_{load}, as shown in Fig. 3. The only parasitic capacitances shown in Fig. 3 are the C_{gd}'s of the p- and n-channel transistors. In this paper, the C_{gd}'s of the p- and n-channel transistors are carefully considered in each region based on the Meyer capacitance model as implemented in SPICE2. Note that C_L is modeled as a voltage-dependent capacitance. C_{gd} is approximated as a piecewise-constant capacitance such that C_{gd} is constant in each region. Accurate modeling of these voltage-dependent capacitances is critical for macromodeling of short-channel devices because input and output transitions are at increasingly higher speeds.

The output node voltage of the inverter, v_o, at a given time is determined by

$$C_{\text{load}}\frac{dv_o}{dt} = I_L = I_p - I_n = I_{dsp} + I_{cgdp} - I_{dsn} + I_{cgdn} \quad (1)$$

where I_L is the current through C_{load}, I_p is the current through the p-channel transistor, and I_n is the current through the n-channel transistor.

For simplicity, the output response here is derived for an input transition from low to high, resulting in a falling output. For the rising input, the contribution of I_p to the change of the output node voltage is negligible while maintaining reasonable accuracy when 1) C_{load} is larger than C_{gd} of the p-channel transistor and the input does not rise fast[1] or 2) the ratio of C_L to the driving capability of the n-channel transistor is large (i.e., the output node voltage pulls down slowly). In such cases of the output falling, I_L can be approximated as $-I_n$ without losing significant accuracy. However, lumping C_{gd} at the output node yields errors for fast input transitions and small output load capacitances. Modeling of this effect is presented later.

The I–V characteristics of an n-channel transistor are modeled in the SPICE level 2 model as

$$I_{dsn} = \beta\frac{\mu_{\text{fact}}}{1-\lambda v_{ds}}\left[\left(v_{gs} - v_{\text{bin}} - \frac{1}{2}\eta v_{dsx}\right)v_{dsx} - \frac{2}{3}\gamma_{\text{SD}}\{(v_{dsx} + PHI - v_{bs})^{3/2} - (PHI - v_{bs})^{3/2}\}\right] \quad (2)$$

where $v_{dsx} = \text{MIN}(v_{ds}, v_{\text{dsat}})$ and v_{dsat} is the saturation voltage, which depends on the operating point. The level 2 model basis provides greater accuracy than can be obtained from the level 1 model. Parameters are defined as follows:

- λ and PHI are SPICE parameters;
- β is $KP(W_{\text{eff}}/L_{\text{eff}})$;
- μ_{fact} is the mobility degradation factor;
- v_{bin} is the effective threshold voltage;
- η is the narrow-width effect parameter;
- γ_{SD} is the effective $GAMMA$ (a SPICE model parameter).

Note that μ_{fact}, v_{bin}, γ_{SD}, v_{dsat}, v_{dsx}, and η are dependent on v_{gs} and v_{ds} while those parameters are constant in the SPICE level 1 model (i.e., secondary effects are ignored). Details of the SPICE level 2 model can be found in [24].

A ramp input voltage to an inverter is represented by

$$v_i = \begin{cases} 0 & \text{for } t \le t_0 \\ \frac{V_{\text{DD}}}{t_r}(t - t_0) & \text{for } t_0 < t \le t_0 + t_r \\ V_{\text{DD}} & \text{for } t > t_0 + t_r \end{cases} \quad (3)$$

where t_r is the rise time of the input waveform, t_0 the time when the input begins to rise, and V_{DD} the power supply voltage. Note that a piecewise-linear input waveform can also be applied in this analysis.

The transient analysis of a CMOS inverter with a typical input slope and output load capacitance is divided into eight regions, as shown in Fig. 4. The motivation and definitions of the various regions are described later in this section. The macromodel analysis consists of 1) estimation of the

[1] "Fast" and "slow" are used here in terms of comparing the gate input slope to the gate output slope. Those are relative terms for each technology.

TABLE II
A SUMMARY OF THE EIGHT REGIONS OF ANALYSIS

region		pmos	nmos	C_L	output modeling	model evaluation
I	$0 < v_i \le v_{tn}$	off	off	–	V_{DD}	no evaluation
II	$v_{tn} < v_i \le v_{inv}$	linear	sat.	–	V_{DD}	no evaluation
III	$v_{inv} < v_i \le v_o - \lvert v_{tp}\rvert$	linear	sat.	C_{load} $+C_{gdp}$	Δv_o for a given Δt, $C_L \frac{dv_o}{dt} \approx \frac{K_{sat}}{1-\lambda v_o}$	
IV	$v_o - \lvert v_{tp}\rvert < v_i$, $v_i \le \text{MIN}(V_{DD} - \lvert v_{tp}\rvert, v_o + v_{tn})$	sat.	sat.	C_{load}	same as in Region III	
V–a	$V_{DD} - \lvert v_{tp}\rvert < v_i < v_o + v_{tn}$	off	sat.	C_{load}	same as in Region III	
V–b	$v_o + v_{tn} < v_i < V_{DD} - \lvert v_{tp}\rvert$	sat.	linear	C_{load} $+C_{gdn}$	same as in Region VI	model evaluation can be avoided
VI	$\text{MAX}(V_{DD} - \lvert v_{tp}\rvert, v_o + v_{tn}) < v_i$, $0.4V_{DD} \le v_o$	off	linear	C_{load} $+C_{gdn}$	$C_L \frac{dv_o}{dt} \approx -I_{dsn}$	model evaluation can be avoided
VII	$V_{DD} - \lvert v_{tp}\rvert < v_i$, $\lvert\Delta v_o\rvert \le v_o < 0.4V_{DD}$	off	linear	C_{load} $+C_{gdn}$	Δt for a given Δv_o	model evaluation can be avoided
VIII	$V_{DD} - \lvert v_{tp}\rvert < v_i$, $0 \le v_o < \lvert\Delta v_o\rvert$	off	linear	C_{load} $+C_{gdn}$	extrapolation of output response	no evaluation (use evaluation from Region VII)

output voltage change over an interval of time; 2) selection of a transistor bias point to estimate the average transistor current value over the interval; 3) computation of the estimated average current value through SPICE model evaluation; and 4) estimation of the actual output voltage change using the average current. This approximation works well as long as input and output voltage changes are small (i.e., less than $0.2V_{\text{DD}}$), which has been implemented here. The six dots represent the specific points for which the modeling routine is evaluated for this circuit. The times at which the MOS model evaluation routine is evaluated is a function of the input and output waveform slopes. Therefore, the model evaluation points vary from circuit to circuit and are not necessarily found in every region. Generally, Regions I and II do not require model evaluations, and neither does Region V in Fig. 4.

The eight regions of the transient analysis can be further simplified. In typical cases when the middle of the output transition occurs after the completion of the input transition, the tail portion (e.g., the last 2 dots in Fig. 4) can be estimated without model evaluation. This is because the transistor operates in the linear region and therefore can be approximated as an effective resistance. The effective resistance, R_{eff}, is approximated using an average of a series of points from dc analysis:

$$R_{\text{eff}} \approx \frac{1}{N}\sum_{i=1}^{N} \frac{v_{dsn,i}}{i_{dsn,i}(v_{gsn,i} = V_{\text{DD}}, v_{dsn,i})} \tag{4}$$

where N is the number of points (e.g., $N = 25$) for $0 < v_{dsn,i} < 0.5V_{\text{DD}}$ using SPICE dc analysis. This dc analysis is done once for a standard size transistor, and the R_{eff} value for each transistor is easily obtained using a scaling factor. Throughout our experiments, the output response of the tail portion, modeled as $v_o(t_{\text{tail}})e^{-(t-t_{\text{tail}})/R_{\text{eff}}C_L}$, produces excellent results except for very slow input transitions.

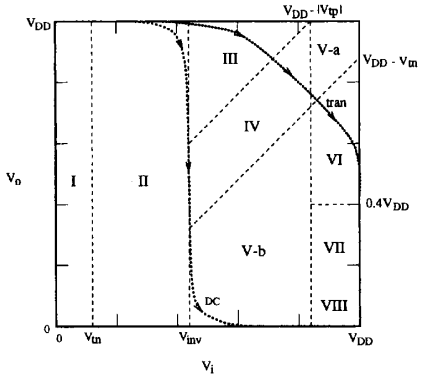


Fig. 5. The regions of operation of the transistors in the v_i versus v_o plane.

Although the simplification of the analysis is possible in most cases, in this paper, the output responses in eight regions are described in order to model very slow input transitions (e.g., when the input slope is 5 times smaller than the output slope) accurately. Fig. 5 illustrates the eight regions of operation of the transistors in the v_i versus v_o plane. These regions are summarized in Table II. Region selections are divided based on operating conditions of the p- and n-channel transistors and v_{inv} (i.e., the unity dc gain point of the inverter). The diagram also shows a typical dc transfer curve for an inverter. In addition, a typical transient curve is shown in the diagram. Typical transitions will pass through regions I, II, III, IV, V-a, VI, VII, and VIII. Gates with relatively slow input transitions as compared to output transitions will be closer to the dc transfer curve and could avoid regions V-a, VI, and VII. Note that the overshoot due to coupling capacitance results in the transient current of Fig. 5 extending above the top of the diagram.

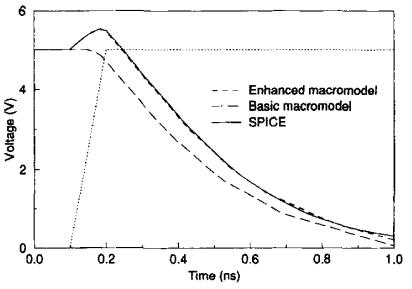


Fig. 6. The effect of coupling capacitances.

In Regions I and II, the output node voltage change is modeled to be zero in typical cases. In Regions III–V, the input voltage is changing and the n-channel transistor is in the saturation region. In these regions the nonlinearity of the n-channel transistor is significant; therefore, approximating the transistor as a linear resistor produces a large delay error. The output responses in these regions are obtained from the solution of $C_L(dv_o/dt) \approx K_{\rm sat}/(1 - \lambda v_o)$, where $K_{\rm sat}$ is defined in Appendix A. Note that there are two possible cases in Region V, depending on operating conditions. The condition in Region V-b occurs rarely (i.e., very small output capacitance or very slow input). In this case, the output response is obtained as in Region VI.

In Regions VI, V-b, VII, and VIII, the n-channel transistor is in the linear region. In typical cases when the input voltage is fixed at $V_{\rm DD}$. the output waveform in these regions can be efficiently estimated using $v_o(t_{\rm tail})e^{-(t-t_{\rm tail})/R_{\rm eff}C_L}$. When the input voltage is changing in these regions, an approximated solution of the differential equation is derived as in Region VI. The output responses in Regions VII and VIII can be obtained using the same solution in Region VI. For efficiency, however, the output response in Region VII is obtained using the technique in ELogic [23] and the output response in Region VIII is extrapolated from the output response in Region VII. Detailed derivations of the output responses are given in the Appendixes.

When the input to an inverter changes several times faster than its output and the output loading is relatively small, the solution using lumped C_{gd}'s at the output node produces significant errors because coupling currents are significant. Fig. 6 shows the resulting error of the basic macromodel in which C_{gd} is lumped at the output node as compared to the result of SPICE2 and the enhanced macromodel presented here, which accounts for C_{gd}'s.

The regions of operation are similarly defined as in Fig. 4, but the output node voltage is obtained differently by modeling the coupling currents through C_{gdp} and C_{gdn}. In this paper, C_{gd} is approximated as a piecewise-constant capacitance such that C_{gd} is constant in each region. Therefore, I_{cgdp} and I_{cgdn} are approximated as

$$I_{cgdp}(t) \approx C_{gdp}(v)\frac{d(v_i - v_o)}{dt} \quad (5)$$

$$I_{cgdn}(t) \approx C_{gdn}(v)\frac{d(v_i - v_o)}{dt}. \quad (6)$$

There are two possible cases in the enhanced macromodel.

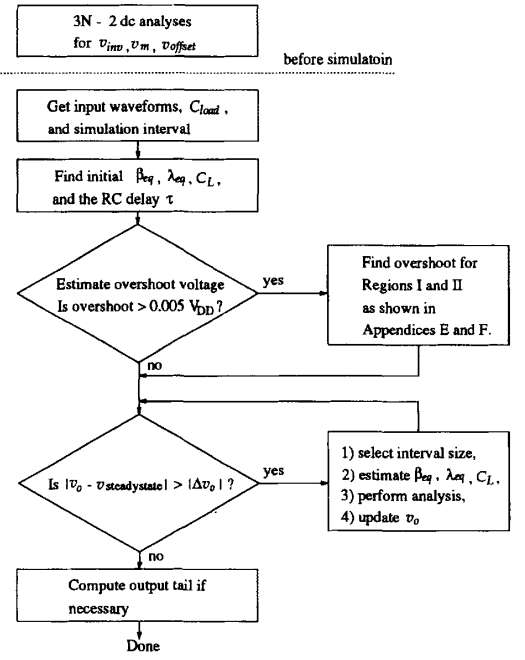


Fig. 7. Simulation flow for processing a gate transition.

Case 1: In this case, the input voltage changes fast and the output load capacitance is small, resulting in an overshoot at the output node such that the resulting output node voltage is still larger than $V_{\rm DD}$ when the input voltage reaches $V_{\rm DD} - |v_{tp}|$, as shown in Fig. 6 (i.e., the p-channel transistor is still active and C_{gdp} does not become zero at the time). The output response of the overshoot portion of the waveform is given in Appendix E. The remainder of the output waveform is computed using the basic macromodel just described.

Case 2: In this case, the input voltage changes relatively fast and/or the output load capacitance is relatively small such that the overshoot is less significant than that of Case 1. If $v_o(t_{\rm inv})$ is less than $V_{\rm DD}$ the basic macromodel of Section II is applied to the remaining regions. If $v_o(t_{\rm inv})$ is larger than $V_{\rm DD}$, $v_o(t_{vtp})$ is estimated as in Case 1. If $v_o(t_{vtp})$ is less than $V_{\rm DD}$, the time when the output waveform crosses $V_{\rm DD}$ is found. Then the remainder of the output waveform is computed using the basic macromodel. The analyses of the overshoot in this case are shown in Appendix F.

III. Reduction of Series-Connected Transistors

Although several nonlinear macromodeling techniques are available, such as IDSIM [13], ILLIADS [14]–[16], and MISIM [17], and that shown in Section II, the reduction of series-connected transistors to the macromodel is a source of significant delay error. This is due to secondary effects: effective channel length modulation, effective transconductance, coupling capacitances, input terminal position dependence, additional delay due to internal capacitances, and the body effect. Parallel-connected transistors can be treated as in [14], [15] and [17]; thus, we concentrate on series configurations here. A simulation flow for a gate is shown in Fig. 7. Estimation of the overshoot is

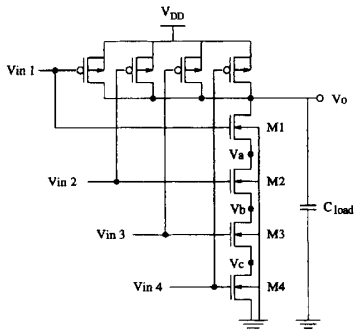


Fig. 8. A 4-input CMOS NAND gate.

given in Appendixes E and F and selection of interval size is discussed in [25].

The following discussion uses a sample 4-input CMOS NAND gate to illustrate the new techniques. In Fig. 8, the gate is shown and the internal voltages are labeled. Appropriate intermediate node voltages under certain conditions are determined using SPICE2 dc analysis of the gate prior to simulation. Since the v_{inv} value for each input position is obtained using dc analysis once, the total number of dc analyses for the v_{inv}'s of N input positions is N (i.e., the number of series-connected transistors). In addition, the calculation of v_m and v_{offset}, described later in this section, for each input position of the series connection (except the bottom input switching) need one time dc analysis. Thus, v_m's and v_{offset}'s require $2(N-1)$ dc analyses. Therefore, total $3N-2$ dc analyses are applied to the gate before simulation to determine steady-state voltages of internal nodes under $3N-2$ operating conditions. These dc analyses are done once for a gate before transient analysis. This calibration helps address many sources of errors in simple primitive macromodel analysis. Estimates of these internal voltages using a simple model will yield increased errors. Note that performing several dc analyses prior to simulation is much faster and simpler than previous techniques which apply many transient analyses. In addition, the $3N-2$ dc analyses applied here are far less than the number required in [26].

The body effect is included in SPICE2 dc analysis of the gate. The body effect of a transistor is often modeled by increasing the threshold voltage of the transistor. Here the body effect (i.e., the decreased current through series-connected transistors due to the increased threshold voltages) is modeled by reducing the effective transconductance of the transistors.

When series-connected transistors are merged into one equivalent transistor in nonlinear macromodeling, the effective channel length modulation, λ_{eq}, has been either ignored, the same as λ [13], or approximated as λ/N [15]. Modeling $\lambda_{\text{eq}} = \lambda/N$ yields better results than ignoring the channel length modulation effect or modeling $\lambda_{\text{eq}} = \lambda$ [15]. However, modeling λ/N is accurate only when inputs are tied together. In this paper, a better approximation of λ_{eq} is presented. The idea comes from acknowledging that the voltage across series-connected transistors is not equally divided to each transistor; the voltage across the topmost transistor can be several times larger than that of the bottom transistor.

In this paper, λ_{eq} is modeled as $((V_{\text{DD}} - v_a)/V_{\text{DD}})\lambda$ in the saturation region instead of λ/N, where v_a is shown in Fig. 8. v_a is determined in dc analysis by setting v_o and all input voltages to V_{DD}. This new λ_{eq} reduces delay errors, particularly when the topmost transistor switches. In the linear region, $\lambda_{\text{eq}} = \lambda/N$ is applied. Note that the delay error due to λ/N during saturation is not negligible when the channel length modulation effect is significant. In addition, this error is counteracted in some cases by errors resulting from considering coupling capacitances inadequately.

The subcircuit output response is a function of the input terminal position, the input slope, parasitic capacitances, and the output load capacitance. Delay dependence on input terminal position is well studied in [26], but not modeled adequately. When the output load capacitance is small and the input transition is fast, the lower terminal switching shows shorter delay. When the output load capacitance is large or the input transition is slow, the upper terminal switching shows shorter delay. This delay dependence on the input terminal position is due to the different transconductance of each transistor and parasitic capacitances in the conducting path.

Traditionally, the effective transconductance, β_{eq}, of N series-connected transistors has been β/N. This approximation is very accurate only when the input is a step waveform, all transistors operate in linear regions, or all inputs of series-connected transistors are tied together. Thus, this is the primary source of delay errors in existing macromodeling techniques. The delay dependence of the input terminal position becomes larger as the input slew rate gets smaller. There are two possible methods of obtaining β_{eq}. The first is estimating v_{gs}, v_{ds}, and v_{bs} of the switching transistor. The other is estimating β_{eq} under various operating conditions based on the results of dc analysis. The latter is chosen because it is easy to incorporate into the primitive macromodel. The modification of β presented here accounts for the adjusted

$$\beta_{\text{eq}} = \begin{cases} \beta\left\{\dfrac{v_{\text{inv}} - v_{\text{offset}}}{v_{\text{inv}}}\dfrac{V_{\text{DD}} - v_m}{V_{\text{DD}}}\right\} = \beta_{\text{eq1}} & \text{for MAX } (v_{in} + v_a, 0.6V_{\text{DD}}) < v_o < V_{\text{DD}} \\ \beta\left\{\dfrac{V_{\text{DD}} - v_{in}}{V_{\text{DD}} - v_{\text{inv}}}\left(\dfrac{\beta_{\text{eq1}}}{\beta} - \alpha_0\right) + \alpha_0\right\} & \text{for } v_{in} - v_{tn} < v_o < \text{MAX } (v_{in} + v_a, 0.6V_{\text{DD}}) \\ \beta\left\{\dfrac{m(V_{\text{DD}} - v_{in})}{V_{\text{DD}} - v_{\text{inv}}}\left(\dfrac{\beta_{\text{eq1}}}{\beta} - \dfrac{1}{N}\right) + \dfrac{1}{N}\right\} & \text{for } 0.4V_{\text{DD}} < v_o < v_{in} - v_{tn} \\ \dfrac{\beta}{N} & \text{for } v_o < 0.4V_{\text{DD}}. \end{cases} \tag{7}$$

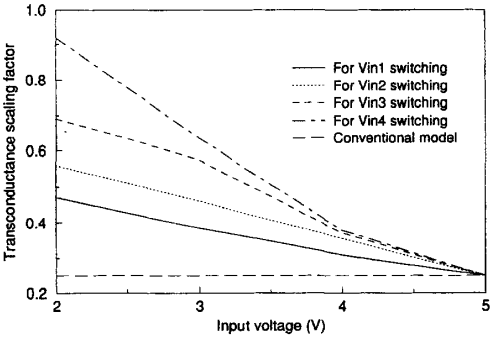


Fig. 9. Typical equivalent β_{eq} scaling factor for macromodeling techniques.

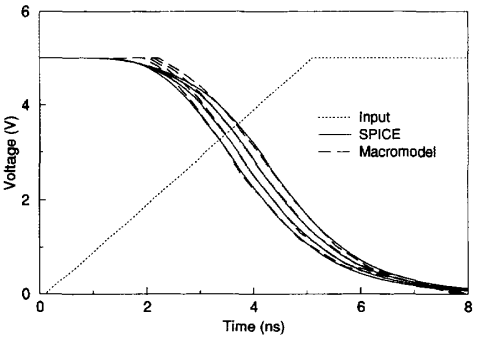


Fig. 10. Macromodeling accuracy using β_{eq}.

bias conditions of the dominant transistor in the series connection. In this paper, a new approximation method based on dc analysis of series-connected transistors is proposed. Basically, β_{eq} varies between β and β/N, depending on operating conditions. The effective scaling of β is obtained using SPICE dc analysis for various input positions of a 4-input CMOS NAND gate, as shown in Fig. 9. The solid line corresponds to the topmost transistor switching. The flat line at 0.25 is the scaling factor applied in existing macromodeling techniques. Note that only v_{gs} values from 2 to 5 V are presented because these are the regions of interest (i.e., β_{eq} need be estimated for only $v_{inv} \leq v_{in} \leq V_{DD}$). β_{eq} is approximated on the previous page. These scaling factors on β account for the reduced v_{ds} and v_{gs} of the dominant transistor in the series connection. $(v_{inv} - v_{offset})/v_{inv}$ accounts for the reduced v_{gs} of the switching transistor. $(V_{DD} - v_m)/V_{DD}$ accounts for the reduced v_{ds} of the switching transistor. The basis for these linear scalings of β due to v_{gs} can be seen in Fig. 9. v_{inv} is the unity dc gain point of the logic gate for the specific input terminal switching, v_m is the value determined in dc analysis at the source node of the switching transistor, v_{offset} is the intermediate node voltage at the source node of the switching transistor when $v_{in} = v_{inv}$. v_{offset} of the nth terminal switching is obtained from dc analysis by setting v_d of the topmost transistor, v_g of the switching transistor, and v_g of the remaining transistors to V_{DD}, v_{inv}, and V_{DD}, respectively. v_m of the nth terminal switching is obtained as v_s of the nth transistor in dc analysis by setting v_d of the nth transistor and v_g of all transistors to V_{DD}. α_0 can be approximated as $1/N$ except for the case when the bottom transistor is switching. $\beta_{eq}(0.4V_{DD} < v_o < v_{in} - v_{tn})$ and the m value are obtained using an average of $\beta_{eq}(v_{in} - v_{tn} < v_o < v_{in} + v_a)$ and $\beta_{eq}(v_o < 0.4V_{DD})$. $m = 0.5$ works well over a wide range of operating conditions.

Fig. 9 shows why this adjustment of β is needed. Delay errors due to the β/N approximation are not as significant if the input slope is much greater than the output slope, the topmost transistor switches, or all inputs are tied together. The accuracy of β_{eq} presented here is demonstrated in Fig. 10. The adjacent pairs of lines are the SPICE2 results and the results of the proposed approximation. It is clear that the position dependence is well accounted for in the macromodel applied in this paper. This has been verified on a wide variety of circuits and input conditions.[2]

This technique for reducing series-connected transistors, in conjunction with existing techniques to process parallel transistor paths [13]–[15], will significantly improve the application of macromodels. Sizing effects and dc voltages are accounted for in the above techniques. The β/N is the equivalent term if the initial β's are the same. The β_{eq} of [13] can be applied for gates with varying device sizes.

Another source of delay error is inadequate consideration of parasitic capacitances, such as coupling capacitances. Four issues must be considered. First, a simple lumping of transistor capacitances is inadequate because the coupling capacitances can significantly contribute to the gate response, as shown in Fig. 6. Therefore, the coupling capacitances must be included in the generic MOS primitive.

Second, the contribution of the coupling capacitance of the topmost transistor in the series connection to the output loading is larger than that of the MOS primitive because the topmost transistor reaches the linear region earlier than the generic MOS primitive. Especially, in contrast to an equivalent primitive transistor, C_{gdn} of the topmost transistor contributes to the output loading during the entire period of output transition except in the case of the top transistor's input switching.

Third, if a coupling capacitance is considered as a portion of device modeling, such as using an equivalent transistor in SPICE2, just determining an effective width, W_{eq}, which may be used to find β_{eq}, results in errors. This is due to the width factor in the capacitance equations. Therefore, it is better to use the transistor width of the switching transistor and scale transconductance to obtain the desired β_{eq}. The inclusion of gate capacitances in the MOS primitive is describe in Section II.

Fourth, the effect of coupling capacitances depends on the input position, as shown in Fig. 2. The capacitances of internal nodes of a gate introduce delay in the gate response. This delay is estimated here based on the capacitances at the internal nodes and β as delay $\tau = (0.75/\beta)(n-1)C_i$. C_i is the internal

[2]In this paper, only single input switching is illustrated. The issue of multiple input switching is addressed in [28] and [29].

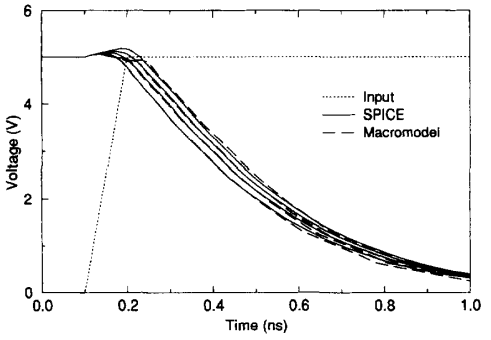


Fig. 11. Macromodeling accuracy considering parasitic capacitances.

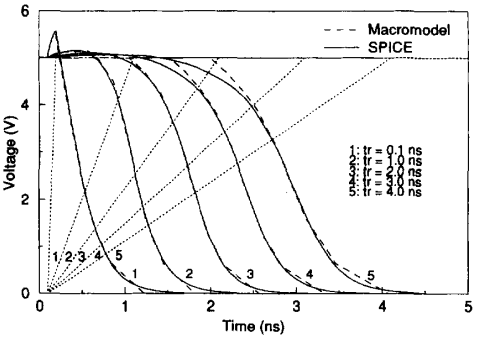


Fig. 12. Simulation results of an inverter for $C_{\text{load}} = 0.05$ pF.

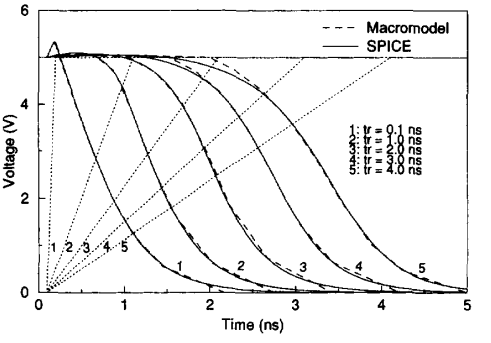


Fig. 13. Simulation results of an inverter for $C_{\text{load}} = 0.1$ pF.

node capacitance and n is 1, 2, 3, and 4 for the topmost, second, third, and bottom transistor switching in the sample 4-input CMOS NAND gate, respectively. The delay estimation is based on the *RC* tree analysis and the common expression of the *RC* delay through a transistor [7]. The delay term due to each specific internal node can be computed individually using β_{eq}'s and appropriate internal node capacitances if the terms are not consistent [7], [27]. The above equation applies when the C_i's are identical. The accuracy of modeling parasitic capacitances presented here is demonstrated in Fig. 11. In [26] I–V characteristics for each switching input are used to estimate β_{eq}. However, because internal node capacitances are not considered in the I–V characteristics, delay errors occur.

Although the reduction techniques proposed are implemented on the nonlinear macromodel based on the SPICE level 2 model, they are generally applicable to other macromodeling techniques, such as in IDSIM2 and ILLIADS. For example, the β_{eq} value for a given input position obtained here can be used in other macromodels instead of β/N. In the special case when all inputs are tied together, which produces the slowest output response, β_{eq} here is β/N. The β_{eq} estimation for series-connected transistors can also be used to estimate the effective resistance of transistors in series in linear macromodels. Similarly, λ_{eff} and coupling capacitance modeling can be easily applied to other macromodeling techniques.

IV. Experimental Results and Discussion

Although many macromodeling techniques are demonstrated through large circuits in typical operating conditions rather than thorough operating region analysis, the validation of the techniques proposed here concentrates on demonstrating the accuracy of nonlinear macromodel and reduction techniques over a wide range of operation of individual gates. This is a more stringent evaluation.

The proposed nonlinear macromodel is validated by applying it to CMOS inverters with a wide variety of input transitions and output loadings (i.e., 0.1 ns $\leq t_r \leq$ 30 ns and $C_{\text{load}} \geq 0.05$ pF). This range of values is based on industrial suggestions [30] (note that the range of input transitions includes extreme cases). Results over such a wide range of operating conditions have not been reported for other macromodeling techniques. The proposed nonlinear macromodel can be applied to larger output loadings (e.g., $C_{\text{load}} > 5$ pF) but errors are expected to be smaller in those cases. The macromodel of Section II produces highly accurate results for larger output loadings. In fact, the larger the output loading capacitance, the simpler and more accurate the analysis.

The output waveforms of a CMOS inverter with $C_{\text{load}} = 0.05$ pF and various input slopes are shown in Fig. 12. The simulation results from SPICE2, the results from the macromodel, and the inputs are plotted by solid lines, dashed lines, and dotted lines, respectively. Notice that the proposed macromodel, which models coupling capacitances, produces much more accurate results than traditional macromodels as shown in Fig. 6.

Figs. 13 and 14 show the output waveforms of inverters to the various input slopes with $C_{\text{load}} = 0.1$ pF and 0.5 pF, respectively. In Fig. 15, the macromodel is applied to the inverter with $C_{\text{load}} = 0.05$ pF and extremely slow input transitions. Our experiments with a wide range of input transitions and output loadings show nearly identical results between SPICE2 and the proposed macromodel. Additional simulation results can be found in [22].

As described before, Figs. 10 and 11 showed the output waveforms of a 4-input CMOS NAND gate for the various input terminals switching with $C_{\text{load}} = 0.5$ pF and $t_r = 5$ ns, and $C_{\text{load}} = 0.1$ pF and $t_r = 0.1$ ns, respectively. Figs. 16 and 17 show the output waveforms of a 4-input CMOS NAND gate for the various input terminals switching with $C_{\text{load}} = 0.2$ pF and $t_r = 2$ ns, and $C_{\text{load}} = 0.5$ pF and $t_r = 3$ ns,

TABLE III
DEVICE MODEL EVALUATION COMPARISON AND SPICE2 EXECUTION TIME FOR AN INVERTER

t_r	.TRAN 0.01 ns 5ns		.TRAN 0.1ns 100ns		.TRAN 0.05ns 5ns		Macromodel	
	# of eval.	Exec. time	# of eval.	Exec. time	# of eval.	Exec. time	# of eval.	Exec. time
0.1 ns	290	1.20 s	321	1.95 s	290	0.68 s	8	0.01 s
1.0 ns	263	1.17 s	317	1.93 s	263	0.65 s	6	0.01 s
2.0 ns	259	1.20 s	313	1.92 s	259	0.65 s	6	0.01 s
3.0 ns	248	1.25 s	292	1.88 s	248	0.67 s	7	0.01 s
4.0 ns	239	1.22 s	282	1.82 s	239	0.67 s	7	0.01 s

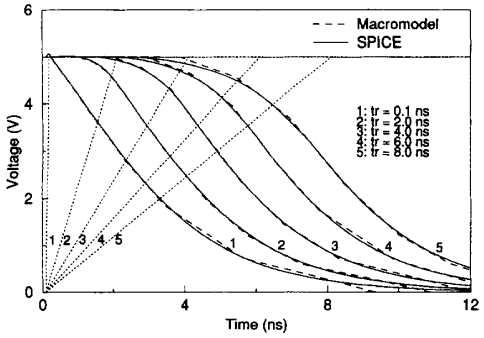


Fig. 14. Simulation results of an inverter for $C_{\text{load}} = 0.5$ pF.

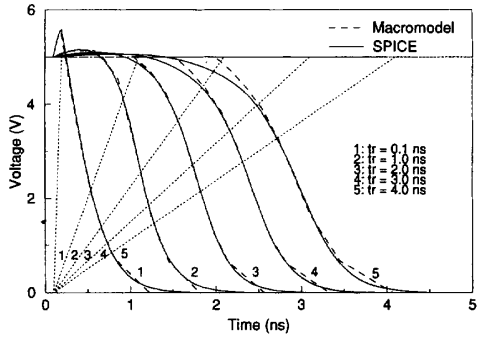


Fig. 15. Simulation results of an inverter for $C_{\text{load}} = 0.05$ pF with slow input slopes.

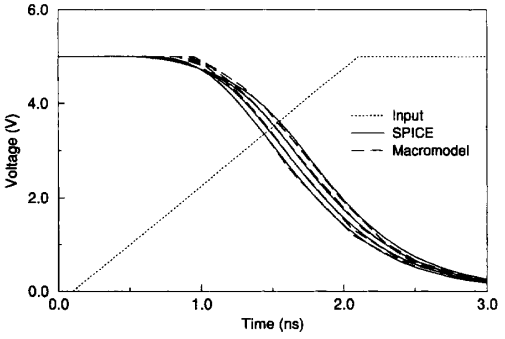


Fig. 16. Simulation results of a 4-input NAND gate for $C_{\text{load}} = 0.2$ pF with various input positions.

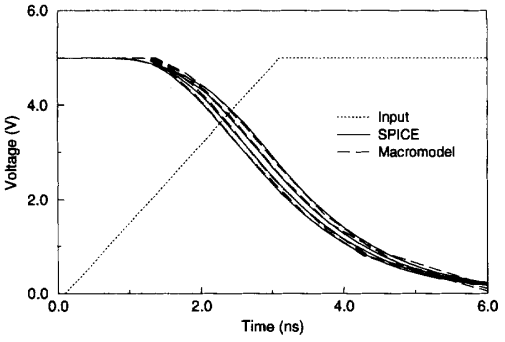


Fig. 17. Simulation results of a 4-input NAND gate for $C_{\text{load}} = 0.5$ pF with various input positions.

respectively. For the above examples, macromodels which apply a single equivalent β and λ have overestimated the delay by up to 68.5% and 43.1%, respectively. The proposed techniques produce very accurate results over a wide range of input slopes, output loadings, and switching input positions.

In most cases, the macromodel evaluates the transistor model equations 5–8 times (4–6 times when $v_o(t_{\text{tail}})e^{-(t-t_{\text{tail}})/R_{\text{eff}}C_L}$ is used) while SPICE2 computes the model at least 230 times for an inverter analysis. For example, the number of model evaluations of the macromodel and SPICE2 in the case of $C_{\text{load}} = 0.05$ pF as shown in Fig. 12 are listed in Table III. The execution times of the macromodel and SPICE2 are also listed in Table III. The simulation time of the macromodel for an inverter is about 0.01 s (i.e., the minimum measurable time on UNIX) while it ranges from 0.6 to several seconds in SPICE2.

The macromodel evaluates the transistor model equations 5–9 times (4–7 times when the exponential tail is used) while SPICE2 computes the model at least 1000 times for a 4-input CMOS NAND gate analysis. For example, the number of model evaluations of the macromodel and SPICE2 in the case of $C_{\text{load}} = 0.5$ pF and $t_r = 3$ ns as shown in Fig. 17 are listed in Table IV. Throughout the experiments, the default error tolerance of SPICE2 is used. The simulation time of the macromodel for a 4-input CMOS NAND gate is still about 0.01 s while it ranges from 1.3 to several seconds in SPICE2.

Our experiments with a wide range of input transitions and output loadings for all combinations of input position of NAND gates as well as inverters show very accurate results as compared to SPICE2. In addition, the proposed techniques are applied to various complex gates, such as AND-OR-INVERT

TABLE IV
DEVICE MODEL EVALUATION COMPARISON AND SPICE2 EXECUTION TIME FOR A 4-INPUT NAND GATE

Input	.TRAN 0.02 ns 10ns		.TRAN 0.1ns 100ns		.TRAN 0.1ns 10ns		Macromodel	
	# of eval.	Exec. time	# of eval.	Exec. time	# of eval.	Exec. time	# of eval.	Exec. time
topmost	1015	1.85 s	1384	2.90 s	1015	1.31 s	6	0.01 s
second	1081	1.90 s	1568	3.05 s	1081	1.42 s	7	0.01 s
third	1035	1.87 s	1501	3.00 s	1035	1.35 s	8	0.01 s
bottom	1058	1.89 s	1527	3.02 s	1058	1.37 s	9	0.01 s

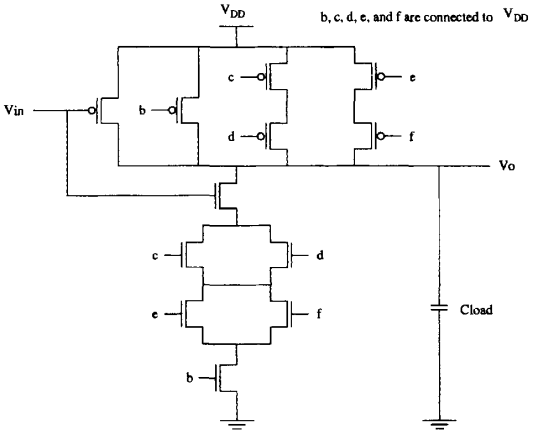


Fig. 18. An AND-OR-INVERT gate with an output load capacitance.

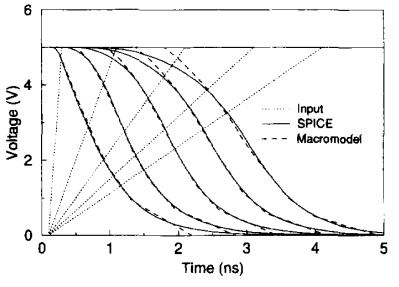


Fig. 19. Input and output waveforms of the AND-OR-INVERT gate using the proposed macromodeling techniques and SPICE2.

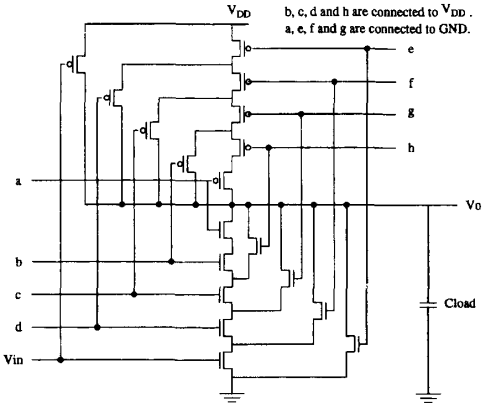


Fig. 20. A carry bit cell with an output load capacitance.

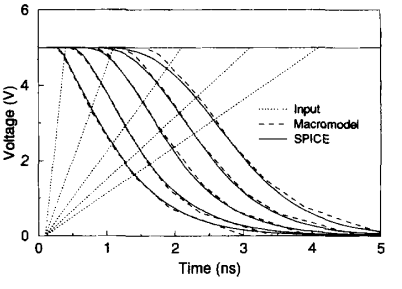


Fig. 21. Input and output waveforms of the carry bit cell using the proposed macromodeling techniques and SPICE2.

gates and carry bit cells in [31]. In Fig. 19, output waveforms of the AND-OR-INVERT gate of Fig. 18 for different input slopes are shown. In Fig. 21, output waveforms of the carry bit cell of Fig. 20 for different input slopes are shown. The results from the macromodel are nearly identical to those of SPICE2. In these cases, the average delay error is less than 2%. The simulation time of the macromodel for the carry bit cell is about 0.01 s while it ranges from 3.3 to 4 s for 100 output data points (i.e., .TRAN 0.05 ns 5 ns) in SPICE2.

These techniques show a simulation speedup of up to several hundred times as compared to SPICE2 and are faster and much more accurate than existing nonlinear macromodels. Throughout the experiments on circuits, including those shown in this paper, transient analyses of the proposed techniques are several times faster than existing nonlinear macromodels, such as IDSIM2 [13] and ILLIADS [14]–[16]. The execution time of the proposed techniques are independent of the TSTEP and TSTOP values because the proposed techniques generate a minimal number of data points required for an accurate output waveform. If the proposed macromodeling techniques are incorporated with a timing simulator, it is expected that the speedup over SPICE2 grows linearly with the circuit size (i.e., at least 3 times the number of transistors in a circuit) [15], [16].

V. CONCLUSIONS

Techniques to account for series-connected transistors in timing macromodeling have been introduced. These techniques address delay errors due to secondary effects, which include

the effective channel length modulation, effective transconductance, coupling capacitances, input terminal dependence, and the body effect. Note that an adequate solution to these sources of errors has not been previously provided. These delay errors, if secondary effects are not addressed adequately in mapping, may total 100% or more.

The techniques to address these secondary effects have been implemented in the nonlinear macromodel described in Section II. Note that the proposed nonlinear macromodel extends analytic macromodeling from the SPICE level 1 to the level 2 model. The application of this nonlinear macromodel to the BSIM model is discussed in [25]. The macromodel is shown to produce accurate results over a wide range of operating conditions. Such a thorough evaluation has not been previously demonstrated. The proposed reduction techniques are not tied to the SPICE level 2 model, but generally applicable to existing nonlinear macromodels.

The results of the application of these new techniques using the described macromodel show that significant improvement in simulation accuracy is obtained. This timing macromodel is up to several hundred times faster than SPICE2 and several times faster than existing nonlinear macromodels for individual gates. This improved speed can be used to minimize VLSI simulation time and provide fast feedback in circuit optimization, such as transistor sizing and transistor reordering [32].

The novel approach of applying several dc analyses to a gate to calibrate the macromodel provides the basis for accurate consideration of various operating conditions. The application of the reduction techniques has little impact on simulation time. The application of calls to model evaluation routines also serves to calibrate the macromodel during simulation. Therefore, the accuracy of existing macromodeling techniques can be significantly improved using these techniques. These techniques also apply to linear and nonlinear macromodels for NMOS and dynamic CMOS circuits.

APPENDIX A
THE OUTPUT RESPONSE IN REGIONS III, IV, AND V

From (1) and (2), the differential equation of the output node voltage is approximated as

$$\begin{aligned} C_L\frac{dv_o(t)}{dt} &\approx -I_n(t) = -I_{dsn}(t) \\ &= -\frac{\beta\mu_{\text{fact}}(t)}{1-\lambda v_o(t)}\Big[(v_i(t)-v_{\text{bin}}(t) \\ &\quad -\frac{1}{2}v_{\text{dsat}}(t))v_{\text{dsat}}(t) \\ &\quad -\frac{2}{3}\gamma_{\text{SD}}(t)\{(v_{\text{dsat}}(t)+PHI)^{3/2}-PHI^{3/2}\}\Big] \\ &\approx -\frac{1}{1-\lambda v_o(t)}(1-\lambda v_o(t_{i-1}+\frac{1}{2}\Delta t_i)) \\ &\quad \cdot I_{dsn}(v_i(t_{i-1}+\frac{1}{2}\Delta t_i), v_o(t_{i-1}+\frac{1}{2}\Delta t_i)) \\ &= -\frac{1}{1-\lambda v_o(t)}K_{\text{sat}} \end{aligned} \tag{8}$$

where C_L is defined in Table II and

$$\begin{aligned} K_{\text{sat}} &= (1-\lambda v_o(t_{i-1}+\frac{1}{2}\Delta t_i)) \\ &\quad \cdot I_{dsn}(v_i(t_{i-1}+\frac{1}{2}\Delta t_i), \\ &\quad v_o(t_{i-1}+\frac{1}{2}\Delta t_i)) \end{aligned} \tag{9}$$

$$\begin{aligned} v_o(t_{i-1}+\frac{1}{2}\Delta t_i) &\approx v_o(t_{i-1}) - \frac{\frac{\Delta t_i}{2}}{0.7R_oC_L}\cdot\frac{V_{\text{DD}}}{2} \\ &= v_o(t_{i-1}) - \frac{\Delta t_i\beta_n V_{\text{DD}}}{2.1C_L}. \end{aligned} \tag{10}$$

By separating variables we have

$$(\lambda v_o(t)-1)dv_o(t) = \frac{K_{\text{sat}}}{C_L}\,dt. \tag{11}$$

By integrating both sides we obtain

$$\lambda v_o^2(t) - 2v_o(t) + 2C_1 = \frac{2K_{\text{sat}}}{C_L}t. \tag{12}$$

The particular solution corresponding to the initial condition at t_{i-1} is

$$\begin{aligned} &\lambda v_o^2(t) - 2v_o(t) + \frac{2K_{\text{sat}}}{C_L}t_{i-1} - \lambda v_o^2(t_{i-1}) + 2v_o(t_{i-1}) \\ &\quad = \frac{2K_{\text{sat}}}{C_L}t. \end{aligned} \tag{13}$$

Therefore,

$$\begin{aligned} &\lambda v_o^2(t_i) - 2v_o(t_i) + \frac{2K_{\text{sat}}}{C_L}t_{i-1} - \lambda v_o^2(t_{i-1}) + 2v_o(t_{i-1}) \\ &\quad = \frac{2K_{\text{sat}}}{C_L}t_i \end{aligned} \tag{14}$$

or

$$\lambda v_o^2(t_i) - 2v_o(t_i) + 2C_{\text{sat}} = 0 \tag{15}$$

where $\Delta t_i = t_i - t_{i-1}$ and

$$C_{\text{sat}} = -\frac{K_{\text{sat}}}{C_L}\Delta t_i - \frac{\lambda}{2}v_o^2(t_{i-1}) + v_o(t_{i-1}). \tag{16}$$

The root, the output response at t_i, is

$$v_o(t_i) = \frac{1-\sqrt{1-2\lambda C_{\text{sat}}}}{\lambda}. \tag{17}$$

The other solution of (15) is not meaningful because λ is generally smaller than 0.2.

APPENDIX B
THE OUTPUT RESPONSE IN REGION VI

The goal of macromodeling in Region VI is to obtain the slope of the line-segment equation, $v_o(t) = B(t-t_{i-1}) + v_o(t_{i-1})$. The differential equation is approximated as

$$\begin{aligned} C_L\frac{dv_o}{dt} &\approx -\frac{\beta\mu_{\text{fact}}}{1-\lambda v_o}\Big[(v_i - v_{\text{bin}} - \frac{1}{2}v_o)v_o \\ &\quad -\frac{2}{3}\gamma_{\text{SD}}\{(v_o+PHI) \end{aligned}$$

$$\cdot\sqrt{v_o(t_{i-1}+\frac{1}{2}\Delta t_i)-|\Delta v_o|+PHI} - PHI^{3/2}\}\Big]. \quad (18)$$

By inserting $v_o(t) = B(t-t_{i-1})+v_o(t_{i-1})$ into (18) we obtain

$$C_L B \approx -\frac{\beta\mu_{\text{fact}}}{1-\lambda(B\Delta t+v_o(t_{i-1}))} \cdot\Big[\{v_i(t)-v_{\text{bin}}-\frac{1}{2}(B\Delta t+v_o(t_{i-1}))\} \cdot(B\Delta t+v_o(t_{i-1}))-\frac{2}{3}\gamma_{\text{SD}}\{(B\Delta t+v_o(t_{i-1}) + PHI)\cdot\sqrt{v_o(t_{i-1})-|\Delta v_o|+PHI}-PHI^{3/2}\}\Big] \quad (19)$$

where $\Delta t = t - t_{i-1}$ and $C_L = C_{\text{load}} + C_{gdn}$.

The slope, B, is obtained from (19) at the middle of the time interval, Δt_i. Let $t = t_{i-1} + (1/2)\Delta t_i$ and $v_i(t) = v_i(t_{i-1} + (1/2)\Delta t_i)$. μ_{fact}, v_{bin}, and γ_{SD} are approximated as functions of $v_i(t_{i-1} + (1/2)\Delta t_i)$ and estimated $v_o(t_{i-1} + (1/2)\Delta t_i)$ as in (10). The default value of $|\Delta v_o|$ is set to $0.2V_{\text{DD}}$ volt. Then multiplying by $2\{1-\lambda(B\Delta t_i/2+v_o(t_{i-1})\}/(\beta\mu_{\text{fact}})$ on both sides and simplification yield

$$\alpha_1 B^2 + 2\alpha_2 B + \alpha_3 = 0 \quad (20)$$

where

$$\alpha_1 = \left(\frac{\Delta t_i}{2}\right)^2 + \frac{C_L\lambda\Delta t_i}{\beta\mu_{\text{fact}}} \quad (21)$$

$$\alpha_2 = \frac{\Delta t_i}{2}\Big\{v_o(t_{i-1}+\frac{1}{2}\Delta t_i) - v_i(t_{i-1}+\frac{1}{2}\Delta t_i) + v_{\text{bin}} + k_1\Big\} - \frac{C_L(1-\lambda v_o(t_{i-1}+\frac{1}{2}\Delta t_i))}{\beta\mu_{\text{fact}}} \quad (22)$$

$$\alpha_3 = v_o^2(t_{i-1}+\frac{1}{2}\Delta t_i) - 2\Big\{v_i(t_{i-1}+\frac{1}{2}\Delta t_i) - v_{\text{bin}} - k_1\Big\}\cdot v_o(t_{i-1}+\frac{1}{2}\Delta t_i) + k_2 \quad (23)$$

$$k_1 = \frac{2}{3}\gamma_{\text{SD}}\sqrt{v_o(t_{i-1}+\frac{1}{2}\Delta t_i)-|\Delta v_o|+PHI}, \quad \text{and} \quad (24)$$

$$k_2 = 2k_1 PHI - \frac{4}{3}\gamma_{\text{SD}} PHI^{3/2}. \quad (25)$$

Thus, the root of (19) is the slope of the output segment from t_{i-1} to t_i as follows:

$$B = \frac{-\alpha_2 - \sqrt{\alpha_2^2 - \alpha_1\alpha_3}}{\alpha_1}. \quad (26)$$

Therefore, the output response at t_i is

$$v_o(t_i) = v_o(t_{i-1}) + B\Delta t_i. \quad (27)$$

APPENDIX C
THE OUTPUT RESPONSE IN REGION VII

The output response can also be obtained from (27) because the inverter operates in the same condition as in Region VI. However, a simpler method is proposed by observing that the output voltage decreases monotonically (or nearly exponentially) in Region VII. Given Δv_o, the corresponding output transition time, Δt, is calculated as in ELogic [23]. The slope of the line segment between $v_o(t_{i-1})$ and $v_o(t_i)$ is approximated as an average of the two derivatives at t_{i-1} and t_i because the output node voltage decreases monotonically. Then the equation of the piecewise-linear segment between $v_o(t_{i-1})$ and $v_o(t_i)$ is

$$v_o(t) = B(t - t_{i-1}) + v_o(t_{i-1}) \quad (28)$$

where

$$B = \frac{\Delta v_o}{\Delta t} \approx \frac{\left\{\frac{dv_o(t)}{dt}\right\}_{t=t_{i-1}} + \left\{\frac{dv_o(t)}{dt}\right\}_{t=t_i}}{2}. \quad (29)$$

Therefore,

$$\Delta t \approx \frac{2\Delta v_o}{\left\{\frac{dv_o(t)}{dt}\right\}_{t=t_{i-1}} + \left\{\frac{dv_o(t)}{dt}\right\}_{t=t_i}} = \frac{-2\Delta v_o C_L}{I_{dsn}(t_{i-1}) + I_{dsn}(t_i, v_o = v_o(t_{i-1}) - |\Delta v_o|)} \quad (30)$$

where Δv_o is given, C_L is $C_{\text{load}}+C_{gdn}$, $I_{dsn}(t_{i-1})$ was calculated at the end of the previous time interval, and $I_{dsn}(t_i, v_o = v_o(t_{i-1}) - |\Delta v_o|)$ can be easily calculated using the MOS model evaluation routine once. Note that the above approximation does not apply to the first time interval in Region VII because $I_{dsn}(t_{i-1})$ was calculated at the middle of the previous interval in Region VI rather than at the end of the previous interval. The above approximation is modified using linear interpolation to be applied at the first time interval and is shown at the bottom of the page. Clearly, this method is simpler than the method used in Region VI where averaging the derivatives does not apply well because the output node voltage does not decrease monotonically. To generate the smoother piecewise-linear output waveform, the last $|\Delta v_o|$ is taken to be a half of $v_o(t_{i-1})$ instead of $0.2V_{\text{DD}}$.

$$\Delta t \approx \frac{(0.5\Delta t_{i-1}+\Delta t_i)\Delta v_o}{0.5\Delta t_{i-1}\cdot\left\{\frac{dv_o(t)}{dt}\right\}_{t=t_{i-1}-0.5\Delta t_{i-1}} + \Delta t_i\cdot\left\{\frac{dv_o(t)}{dt}\right\}_{t=t_i}} \approx \frac{-3\Delta v_o C_L}{I_{dsn}(t_{i-1}-0.5\Delta t_{i-1}) + 2I_{dsn}(t_i, v_o = v_o(t_{i-1}) - |\Delta v_o|)}. \quad (31)$$

APPENDIX D
THE OUTPUT RESPONSE IN REGION VIII

The output response is approximated as a quadratic equation, $v_o(t) = a_1(t-b_1)^2$. The output response can be approximated as an exponential function, $v_o(t) = V_{DD}e^{-a_2(t-b_2)}$, as well. The first output point in this region is calculated from the previous analysis in Region VII. The output response at the t_{stop} or the time when the output node voltage reaches zero is approximated without model evaluation.

The output response at t_i is following:

$$v_o(t_i) = \begin{cases} v_o(t_{stop}) = a_1(t_{stop}-b_1)^2 & \text{if } t_{stop} < b_1 \\ v_o(b_1) = v_o(t_{stop}) = 0 & \text{otherwise} \end{cases} \quad (32)$$

where t_{stop} is the time of the last transient analysis,

$$a_1 = \frac{v_o(t_{i-2})}{(t_{i-2}-b_1)^2}, \quad b_1 = \frac{t_{i-1}r_1 - t_{i-2}}{r_1 - 1}, \quad \text{and}$$
$$r_1 = \sqrt{\frac{v_o(t_{i-2})}{v_o(t_{i-1})}}. \quad (33)$$

APPENDIX E
THE OUTPUT RESPONSE IN CASE 1 OF SECTION II

The output responses are following.

$$v_o(t_{vtn}) \approx V_{DD}\left\{1 + \frac{C_{gdp}(t_{vtn}-t_0)}{C_L t_r}\right\}$$

$$v_o(t_{inv}) \approx v_o(t_{vtn}) + \frac{V_{DD}C_{gdp}(t_{inv}-t_{vtn})}{C_L t_r} \quad (34)$$
$$- \frac{I_{dsn}(t_{inv})(t_{inv}-t_{vtn})}{2C_L} \quad (35)$$

$$v_o(t_{vtp}) \approx v_o(t_{inv}) + \frac{V_{DD}C_{gdp}(t_{vtp}-t_{inv})}{C_L t_r}$$
$$- \frac{(I_{dsn}(t_{inv}) + I_{dsn}(t_{vtp}))(t_{vtp}-t_{inv})}{2C_L} \quad (36)$$

$$v_o(t_0+t_r) \approx v_o(t_{vtp}) + \frac{V_{DD}C_{gdp}(t_0+t_r-t_{vtp})}{2C_L t_r}$$
$$- \left\{ \frac{(I_{dsn}(t_{vtp}) - I_{dsn}(t_{inv}))\left(\frac{t_0+t_r+t_{vtp}}{2} - t_{inv}\right)}{t_{vtp}-t_{inv}} + I_{dsn}(t_{inv}) \right\} \frac{t_0+t_r-t_{vtp}}{C_{load}} \quad (37)$$

where C_L is $C_{load} + C_{gdp}$, t_{vtn} is the time when $v_i = v_{tn}$, t_0 is the time when input begins to rise, t_r is the rise time of the input waveform (from 0 to V_{DD}), t_{inv} is the time at the unity dc gain point of the inverter, t_{vtp} is the time when $v_i = v_{tp}$, $I_{dsn}(t_{inv}) \approx I_{dsn}(v_o = V_{DD}, v_i = v_{inv})$, and $I_{dsn}(t_{vtp}) \approx I_{dsn}(v_o = V_{DD}, v_i = v_{vtp})$. The output response in other regions is obtained in the same procedure as in the basic macromodel. $v_o(t_{vtn})$ and $v_o(t_{vtp})$ terms are used as the overshoot estimate (e.g., $v_o(t_{vtn}) > 1.002V_{DD}$ and $v_o(t_{vtp}) > 1.005V_{DD}$).

APPENDIX F
THE DERIVATION OF $t(v_o = V_{DD})$ IN CASE 2 OF SECTION II

When $v_{inv} < v_i < V_{DD} - |v_{tp}|$, I_{dsn} is approximated as a piecewise-linear function of time and the resulting differential equation at the output node is approximated as

$$C_{load}\frac{dv_o(t)}{dt} \approx I_{cgdp}(t) - I_{dsn}(t) \approx C_{gdp}\frac{d(v_i(t)-v_o(t))}{dt} - \{z_1(t-t_{inv}) + I_{dsn}(t_{inv})\} \quad (38)$$

or

$$C_L\frac{dv_o(t)}{dt} \approx \frac{C_{gdp}V_{DD}}{t_r} - I_{dsn}(t_{inv}) - z_1(t-t_{inv}) \quad (39)$$

where $C_L = C_{load} + C_{cgdp}$ and

$$z_1 = \frac{I_{dsn}(t_{vtp}) - I_{dsn}(t_{inv})}{t_{vtp}-t_{inv}}. \quad (40)$$

Integrating (39) yields

$$C_L v_o(t) \approx \left(\frac{C_{gdp}V_{DD}}{t_r} + z_1 t_{inv} - I_{dsn}(t_{inv})\right)t - \frac{z_1}{2}t^2 + C_1. \quad (41)$$

The particular solution corresponding to the initial condition at t_{inv} is given by

$$C_1 = C_L v_o(t_{inv}) - \left(\frac{C_{gdp}V_{DD}}{t_r} + z_1 t_{inv} - I_{dsn}(t_{inv})\right) \cdot t_{inv} + \frac{z_1}{2}t_{inv}^2. \quad (42)$$

The time when the output waveform crosses V_{DD} is obtained by solving $V_{DD} = v_o(t)$ in (41). By rearranging $V_{DD} = v_o(t)$, we obtain

$$t^2 - 2z_2 t - z_3 = 0 \quad (43)$$

where

$$z_2 = \frac{1}{z_1}\left(\frac{C_{gdp}V_{DD}}{t_r} + z_1 t_{inv} - I_{dsn}(t_{inv})\right) \quad (44)$$
$$z_3 = \frac{2}{z_1}(C_1 - C_L V_{DD}). \quad (45)$$

Therefore, the time when $v_o(t) = V_{DD}$ is

$$t(v_o = V_{DD}) = z_2 + \sqrt{z_2^2 + z_3}. \quad (46)$$

After $t(v_o = V_{DD})$, the basic macromodel is applied in the remaining regions. Note that the basic macromodel for Case 2 models I_{cgdp} as in (5) for $v_i < V_{DD} - |v_{tp}|$.

ACKNOWLEDGMENT

The authors wish to thank Dr. K. H. Kim of Samsung Electronics Co., Korea, and Prof. S. M. Kang and Dr. A. Dharchoudhury of the University of Illinois at Urbana for valuable discussions. The authors would also like to thank the reviewers for their constructive suggestions.

REFERENCES

[1] L. W. Nagel, "SPICE2: A computer program to simulate semiconductor circuits," Electron. Res. Lab., Univ. of California, Berkeley, Memo. ERL-M520, May 1975.

[2] J.-T. Kong and D. Overhauser, "A survey of simulation and macromodeling techniques for VLSI circuits," Dep. of Electrical Eng., Duke Univ., Durham, NC, Tech. Rep. DUKE-CCSR-92-010, Nov. 1992.

[3] L. T. Pillage and R. A. Rohrer, "Asymptotic waveform evaluation for timing analysis," *IEEE Trans. Computer-Aided Design*, vol. 9, no. 4, pp. 352–366, Apr. 1990.

[4] S. M. Kang, "A design of CMOS polycells for LSI circuits," *IEEE Trans. Circuits Syst.*, vol. CAS-28, no. 8, pp. 838–843, Aug. 1981.

[5] N. Weste and K. Eshraghian, *Principles of CMOS VLSI Design*, 2nd ed. Reading, MA: Addison–Wesley, 1993.

[6] D. A. Hodges and H. G. Jackson, *Analysis and Design of Digital Integrated Circuits*, 2nd ed. New York: McGraw–Hill, 1988.

[7] J. P. Uyemura, *Circuit Design for CMOS VLSI*. Norwell, MA: Kluwer, 1992.

[8] J. R. Burns, "Switching response of complementary-symmetry MOS transistor logic circuits," *RCA Rev.*, vol. XXV, no. 4, pp. 627–661, Dec. 1964.

[9] N. Hedenstierna and K. O. Jeppson, "CMOS circuit speed and buffer optimization," *IEEE Trans. Computer-Aided Design*, vol. CAD-6, no. 2, pp. 270–281, Mar. 1987.

[10] K. H. Kim and S. B. Park, "Delay-time modelling and critical-path verification for CMOS digital designs," *Computer-Aided Design*, vol. 23, no. 9, pp. 604–614, Nov. 1991.

[11] Y.-H. Jun and I. N. Hajj, "An efficient timing simulation approach for CMOS digital circuits," in *Proc. IEEE Midwest Symp. Circuits Syst.*, May 1990, pp. 235–238.

[12] K. O. Jeppson, "Modeling the influence of the transistor gain ratio and the input-to-output coupling capacitance on the CMOS inverter delay," *IEEE J. Solid-State Circuits*, vol. 29, no. 6, pp. 646–654, June 1994.

[13] D. V. Overhauser, "Fast timing simulation of MOS VLSI circuits," Ph.D. dissertation, Dep. of Electrical Eng., Univ. of Illinois, Urbana, Aug. 1989.

[14] Y.-H. Shih and S.-M. Kang, "Analytic transient solution of general MOS circuit primitives," *IEEE Trans. Computer-Aided Design*, vol. 11, no. 6, pp. 719–731, June 1992.

[15] Y.-H. Shih, "Computationally efficient methods for accurate timing and reliability simulation of ultra-large MOS circuits," Ph.D. dissertation, Dep. of Comput. Sci., Univ. of Illinois, Urbana, July 1991.

[16] Y.-H. Shih, Y. Leblebici, and S.-M. Kang, "ILLIADS: A fast timing and reliability simulator for digital MOS circuits," *IEEE Trans. Computer-Aided Design*, vol. 11, no. 6, pp. 1387–1402, Sept. 1993.

[17] Y.-H. Chang and A. T. Yang, "Analytic macromodeling and simulation of tightly-coupled mixed analog-digital circuits," in *Proc. IEEE Int. Conf. Computer-Aided Design*, Nov. 1992, pp. 244–247.

[18] T. Sakurai and A. R. Newton, "A simple MOSFET model for circuit analysis," *IEEE Trans. Electron Devices*, vol. 38, no. 4, pp. 887–893, Apr. 1991.

[19] S. M. Kang and A. Dharchoudhury, private communication, June 1993.

[20] N. B. Rabbat, W. D. Ryan, and S. Q. A. M. A. Hossain, "A computer modeling approach for LSI digital structures," *IEEE Trans. Electron Devices*, vol. ED-22, no. 8, pp. 523–531, Aug. 1975.

[21] A. H. M. Shousha and M. Aboulwafa, "A generalized tanh law MOSFET model and its applications to CMOS inverters," *IEEE J. Solid-State Circuits*, vol. 28, no. 2, pp. 176–179, Feb. 1993.

[22] J.-T. Kong and D. Overhauser, "Timing macromodeling for submicron CMOS," Dep. of Electrical Eng., Duke Univ., Tech. Rep. DUKE-CCSR-93-008, Apr. 1993.

[23] Y. H. Kim, S. H. Hwang, and A. R. Newton, "Electrical-logic simulation and its applications," *IEEE Trans. Computer-Aided Design*, vol. 8, no. 1, pp. 8–22, Jan. 1989.

[24] D. A. Divekar, *FET Modeling for Circuit Simulation*. Norwell, MA: Kluwer, 1988.

[25] J.-T. Kong, "Fast-timing macromodeling for VLSI design verification," Ph.D. dissertation, Dep. of Electrical and Comput. Eng., Duke Univ., Durham, NC, Nov. 1994.

[26] T. Sakurai and A. R. Newton, "Delay analysis of series-connected MOSFET circuits," *IEEE J. Solid-State Circuits*, vol. 26, no. 2, pp. 122–131, Feb. 1991.

[27] L. Wurtz, "An efficient scaling procedure for domino CMOS logic," *IEEE J. Solid-State Circuits*, vol. 28, no. 9, pp. 979–982, Sept. 1993.

[28] Y.-H. Jun, K. Jun, and S.-B. Park, "An accurate and efficient delay time modeling for MOS logic circuits using polynomial approximation," *IEEE Trans. Computer-Aided Design*, vol. 8, no. 9, pp. 1027–1032, Sept. 1989.

[29] A. Nabavi-Lishi and N. C. Rumin, "Inverter models of CMOS gates for supply current and delay estimation," *IEEE Trans. Computer-Aided Design*, vol. 13, no. 10, pp. 1271–1279, Oct. 1994.

[30] K. H. Kim at Samsung Electronics Co., private communication, Jan. 1993.

[31] D. Deschacht, M. Robert, and D. Auvergne, "Synchronous-mode evaluation of delays in CMOS structures," *IEEE J. Solid-State Circuits*, vol. 26, no. 5, pp. 789–795, May 1991.

[32] B. S. Carlson and C. Y. R. Chen, "Performance enhancement of CMOS VLSI circuits by transistor reordering," in *Proc. ACM/IEEE Design Automat. Conf.*, June 1993, pp. 361–366.

Jeong-Taek Kong received the B.S. degree in electronics engineering from Hanyang University, Korea, in 1981, the M.S. degree in electronics engineering from Yonsei University, Korea, in 1983, and the Ph.D. degree from Duke University, Durham, NC, in 1994.

From 1983 to 1990, he was with Samsung Electronics Co., Korea, as a VLSI CAD Manager. From 1990 to 1994, he was at Duke University granted by a Fellowship from Samsung Electronics Co., Ltd. Currently, he is with Semiconductor Business, Samsung Electronics Co., Korea, as a CAD Manager. His research interests focus on various VLSI CAD tools.

Dr. Kong has authored and coauthored several papers in the VLSI CAD area and coauthored a book entitled *Digital Timing Macromodeling for VLSI Design Verification*.

David Overhauser (S'86–M'89) received the B.S. degree in math and computer science from Purdue University, West Lafayette, IN, in 1983 and the M.S. and Ph.D. degrees from the University of Illinois, Urbana-Champaign, in 1985 and 1989, respectively.

In 1989 he joined the Department of Electrical Engineering at Duke University. He is currently engaged in teaching and research pertaining to VLSI design, VLSI CAD tool design, circuit simulation, and mixed-mode simulation.

ELSEVIER

Microelectronic Engineering 31 (1996) 319-330

MICROELECTRONIC ENGINEERING

VLSI Testing with CAD-Linked Electron Beam Test System

Koji NAKAMAE, Katsuyoshi MIURA and Hiromu FUJIOKA[a]

[a]Department of Information Systems Engineering, Faculty of Engineering, Osaka University, Yamada-Oka 2-1, Suita, Osaka 565, Japan

Two approaches for efficient fault diagnosis are proposed where only CAD layout data is available in the CAD-linked electron beam test system. One is a circuit logical function extraction method from the CAD layout data. In the method, after extracting transistor-level circuit data in a unit of a primitive logic gate from a flat structured CAD layout data, the data are translated to higher cell-level data automatically in order to trace a fault hierarchically. The other is a hierarchical fault tracing method for a hierarchically structured CAD layout data. The method allows us to trace a fault hierarchically from the top level cell to the lowest primitive cell and from the primitive cell to the transistor-level circuit in a consistent manner independent of circuit functions even when the cell data and the transistor-level circuit data exist in a level as a mixture.

1. INTRODUCTION

As LSI circuits have become denser and more complex, the performance faults such as the delay fault have been a serious issue in the diagnosis of faulty VLSI chips [1]. The CAD-linked electron-beam (EB) test system has become an indispensable instrument for probing internal behavior of VLSI circuits [2]. Through the fault localization process by using the method such as the guided probe diagnosis with the CAD-linked EB test system, the gate or cell that has the faulty output and all good inputs is found: a gate-level fault or a cell-level fault is determined [3], [4], [5]. In order to search the cause of the performance faults, however, it becomes necessary to trace the fault in the lowest level of the circuits, that is, in the transistor-level [6].

In some situations of the CAD database system, there often arises the case where the transistor-level data is lost and only a layout data is available [7]. For example, after the design process is completed, the intermediate data may be disposed in order to save the storage capacity, keeping only final data such as a layout data. In regard to the fault diagnosis of ASICs where the design does not always require the transistor-level data, a gate-level netlist is used [8]. In those cases, it is necessary to extract the transistor-level netlist from the layout data.

We have recently proposed an automatic transistor-level performance fault tracing method [9] which is applicable to the case where only CAD layout data is available. The method uses an integrated algorithm that combines a previously proposed transistor-level fault tracing algorithm [6] and a successive circuit extraction from a non-hierarchically or a flat structured CAD layout data. This enables us to trace a fault in the transistor level by using only CAD layout data. If the circuit logical function is identified in the method, the efficient fault tracing might be possible by carrying out the transistor-level fault tracing after tracing the fault in the gate-level or the

SSDI 0167-9317(95)00354-1

cell-level.

A number of VLSI checking programs have been developed for verifying the geometrical layout data to determine whether or not it correctly reflects the original logic level description [10], [11], [12]. In these programs, after having extracted the transistor-level netlist, the netlist is compared with a reference netlist hierarchically and discrepancies are indicated. For the purpose of fault tracing, however, it is sufficient to extract partial data along the tracing path from the layout data and to translate them to higher level circuit data. Furthermore, it is convenient in a hierarchical fault tracing to extract circuit data in a unit of a primitive logic gate such as an inverter, a NAND gate or a NOR gate and so on.

A modern complex VLSI system is generally composed of subsystems which can be structured separately. For such a system, the hierarchical layout structure is the most general [13], [14]. This motivates us to develop a hierarchical approach by utilizing the hierarchical layout structure.

In this paper, we propose two approaches for efficient fault diagnosis where only CAD layout data is available in the CAD-linked electron beam test system. One is a circuit logical function extraction method from the CAD layout data. In the method, after extracting transistor-level circuit data in a unit of a primitive logic gate from a flat structured CAD layout data, the data are translated to higher cell-level data automatically. The other is a hierarchical fault tracing method by utilizing a hierarchical layout structure. The method allows us to trace a fault hierarchically from the top level cell to the lowest primitive cell and from the primitive cell to the transistor-level circuit in a consistent manner independent of circuit functions. Both methods use the expansion of a previously proposed integrated algorithm [9] that combines a transistor-level fault tracing algorithm [6] and a successive circuit extraction from a flat structured CAD layout data. In the next section, the outline of automatic transistor-level fault tracing by successive circuit extraction is described. The circuit logical function extraction method for a flat structured CAD layout data and the hierarchical fault tracing method for a hierarchically structured CAD layout data are described in sections 3 and 4, respectively. The applied results are presented in section 5.

2. OUTLINE OF THE FAULT TRACING BY SUCCESSIVE CIRCUIT EXTRACTION FROM FLAT STRUCTURED CAD LAYOUT

We briefly review the fault tracing by successive circuit extraction from a flat structured CAD layout data [9]. In the method, the transistor-level fault tracing algorithm [6] and a successive circuit extraction from a flat structured CAD layout data are integrated.

2.1. Transistor-level fault tracing algorithm

We model the problems in the transistor-level automatic fault tracing of MOS combinational circuits using a graph where a single fault is assumed. The circuit is translated into a circuit connection graph where the interconnection is a vertex and the MOS transistor is an edge shown by a solid line which is associated with a dotted line corresponding to the gate electrode. As an example, the circuit connection graph translated from a CMOS circuit shown in Fig. 1 is illustrated in Fig. 2. In these figures, numbers are put on the interconnections and small letters on the MOS transistors.

Our algorithm utilizes *Depth-First Search* algorithm where graph is searched as deeply as possible. In the algorithm, we use the labels called *DC PATH LABEL*s that are put on the

interconnections and the devices which constitute a DC path (an ELEMENTARY PATH). A DC path is defined as the path which connects the starting vertex to the power supplies V_{DD} and V_{SS}.

In Figs. 1 and 2, it is assumed that the interconnection labeled 1 has a fault and a faulty signal propagates along interconnections drawn by thick solid lines. We start a fault tracing at the interconnection 8. As shown in Figs. 1 and 2, we have two trace paths at the interconnection 2. For fault tracing to the upstream, the path to the vertices labeled 3, 4 and 7 should be ignored. The upstream vertices are searched along the DC path which connects the interconnection 8 to the power supplies V_{DD} and V_{SS}: *DC PATH LABELs* are put on the interconnections 8 and 2, and the MOS transistors e, f, a and b with the gate terminals 5, 6 and 1. In the algorithm, an upstream vertex is selected for further exploration from the most recently visited vertex. Then the waveform on the vertex is measured by an EB tester and is compared with the expected data. If the signal is determined faulty through the comparison, a next upstream vertex is selected. In Figs. 1 and 2, the gate electrode 5 or 6 and the interconnections 2 and 1 are tested. The procedures described above are repeated until good signals are obtained on all upstream vertices. If such vertices are found, it is determined that the origin of the fault is localized between the present visited and the most recently visited vertices.

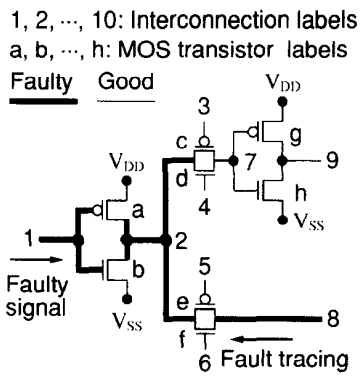


Figure 1. An example of CMOS circuit. Numbers are put on interconnections and small letters on MOS transistors. The thick solid lines indicate a faulty signal propagation path.

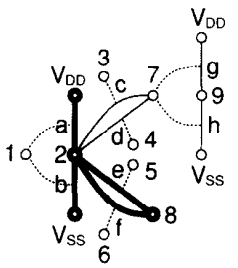


Figure 2. Circuit connection graph translated from the circuit shown in Fig. 1.

2.2. Successive circuit extraction from flat structured layout data

The fault tracing by successive circuit extraction from a flat structured CAD layout data consists of the pre-processing of the inputted layout data and the fault tracing by successive extraction of circuit connectivity information. CMOS circuits with a flat structured CAD layout data are assumed. The extraction method operates only on Manhattan layouts, though the method can easily be expanded so as to handle arbitrary angle trapezoidal layouts.

In order to execute extraction processes efficiently, two types of data structures are used as shown in Fig. 3: a circuit data structure and a quad tree data structure. The two structures are linked each other by links labeled ℓ_1 and ℓ_3. The circuit data structure consists of interconnection structures and device structures, and the mutual link labeled ℓ_2. The quad tree is an organization of rectangular area that contains some rectangles in a hierarchical fashion, and enables us to search polygons quickly.

Interconnection structures and the quad tree data structure, and their mutual link ℓ_1 are built in the pre-processing, while device structures in the circuit data structure and the links ℓ_2 and ℓ_3

are formed in the successive circuit data extraction process.

In the pre-processing, various processing such as forming "device" layer which consists of polygons (device layer polygons) representing MOS transistor channel regions and "SD" layer which consists of polygons (SD layer polygons) representing source/drain regions, constructing quad tree data structure which enables us to search polygons quickly, extracting interconnections to build the interconnection structures linked with the quad tree data structure, and indicating the interconnections for power supplies are carried out.

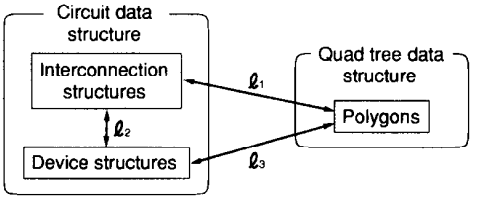


Figure 3. Data structure used in the non-hierarchical fault tracing.

In the fault tracing by successive extraction, the constituent polygons of a start interconnection are known from the link of the interconnection structure to the quad tree data structure. Device polygons connected to the interconnection constituent polygons are searched in the quad tree data structure. The device structures linked to the quad tree data structure are built and are further linked to the interconnection structure. Then, SD polygons connected to the device polygons are searched in the quad tree data structure. From the links of these polygons to the interconnection structures, the interconnections connected to the devices are known and are linked to the device structures. These processes are repeated until the procedure reaches the interconnections for power supplies V_{DD} and V_{SS}. In this way, a path to the power supply lines is obtained: a DC path is constructed. The transistor-level fault tracing algorithm selects a gate terminal of one transistor. A polygon which represents the gate terminal is searched in the quad tree data structure. From the link of the polygon to the interconnection structure, the gate interconnection is obtained. The signal waveform on the interconnection is measured by the EB tester and is compared with the reference waveform. When the signal is faulty, the procedure is repeated with the gate interconnection as a start point until all upstream interconnections representing good signal waveforms are found.

3. CIRCUIT LOGICAL FUNCTION EXTRACTION FOR FLAT STRUCTURED CAD LAYOUT DATA

A circuit logical function extraction from the CAD layout data for a hierarchical fault tracing consists of two steps: 1) transistor-level circuit data extraction from the CAD layout data by utilizing the successive circuit extraction algorithm, and 2) automatic circuit data translation of transistor-level data to higher cell-level data hierarchically.

3.1. Transistor-level circuit data extraction

The circuit data extraction consists of the pre-processing of the inputted layout data and the extraction of circuit connectivity information (transistor-level circuit data). The pre-processing and the data structures are the same as those described in section 2.2.

The extraction of transistor-level circuit data is carried out in a unit of a primitive logic gate along the path from the start interconnection to the upstream signal flow direction by utilizing *Depth-First Search* algorithm. A primitive logic gate is defined as a gate which constitutes only one DC path in the transistor-level circuit such as an inverter, a NAND gate or a NOR gate.

Based on this definition, an AND gate is composed of two primitive logic gates, a NAND gate and an inverter. The transmission gate cannot be a primitive logic gate, since the transmission gate does not make up a DC path. In the circuit which includes the transmission gates, the sum of the transmission gate and the upstream primitive logic gate is extracted as one unit. When there exists plural paths from the start interconnection to some primitive logic gate in the upstream signal flow direction, the primitive logic gate is processed along the shortest path.

The circuit data extraction outputs a circuit data in a SPICE-like network description.

3.2. Circuit data translation

In the translation of transistor-level circuit data to the higher level circuit data hierarchically, we adapt a hybrid approach which combines the logical synthesis technique [10] and the library-based graph-isomorphism technique [12].

In the logical synthesis process [10], the primitive logical circuit data extracted on a DC path is translated into a graph G which represents transistor connections. The graph G is divided into two subgraphs H_n and H_p where only nMOS or pMOS transistors are included respectively. Both subgraphs are reduced by series reduction and parallel reduction until each reduced graph consists of one edge. The logical synthesis technique allows us to recognize primitive logic gates such as inverter, NAND, NOR, AND-OR-INVERTER and OR-AND-INVERTER gates.

In the library-based graph-isomorphism technique [12], the template matching is carried out where a template is compared to all unknown subgraph in the target graph. The template is the graph which corresponds to a reference netlist in a cell library. The target graph corresponds to the netlist extracted from the CAD layout or the netlist after the logical synthesis is performed. The library-based graph-isomorphism technique allows us to recognize logic gates or cells such as transmission gate, tristate buffer, AND, OR, XOR, XNOR, multiplexer, half adder, full adder and D flip flop at present.

The hybrid approach consists of three stages. First stage is the identification of pass transistors by the library-based graph-isomorphism technique. In the second stage, primitive logic gates are identified by the logical synthesis technique. In the final stage, higher level cells are recognized hierarchically by template matching in the library-based graph-isomorphism technique.

The method gives us the netlist where the net name is unique in spite of the level in hierarchy.

4. HIERARCHICAL FAULT TRACING FOR HIERARCHICALLY STRUCTURED CAD LAYOUT DATA

In a hierarchical fault tracing method by utilizing a hierarchical layout structure, the previous integrated algorithm for a flat structured CAD layout data [9] is expanded so as to allow us to trace a fault hierarchically from the top level cell to the lowest primitive cell and from the primitive cell to the transistor-level circuit in a consistent manner even when the cell data and the transistor-level circuit data are mixed in a level.

4.1. Hierarchically structured CAD layout data

A modern complex VLSI chip is generally structured by a hierarchy of cells where the design is divided along functional boundaries into smaller pieces called cells and those cells are subsequently divided into even smaller cells until the problem is manageable [13]. Each cell is composed of primitive objects (such as polygons) and references to other cells called instances. The cell instance is treated as a bounding box which is the smallest rectangle that includes all

layout in the cell. The top level cell corresponds to the whole design of a chip.

On a hierarchically structured CAD layout data of MOS combinational circuits, following assumptions are added to the assumptions on a flat structured layout data.

- The lowest level primitive cell is the primitive logic gate such as an inverter.
- Cells in a level do not overlap and abut each other.
- Interconnections do not pass through the cells.
- There is no bi-directional interconnections.

4.2. Hierarchical fault tracing by utilizing hierarchical layout structure

The hierarchical fault tracing consists of the general pre-processing of the inputted layout data, the successive pre-processing of a cell layout data and the fault tracing by successive extraction of circuit connectivity information.

4.2.1. General pre-processing

In the general pre-processing, the data structure called a cell structure as illustrated in Fig. 4 is built for each cell included in the whole layout and the interconnections for power supplies are indicated. A cell structure consists of a circuit data structure and a quad tree data structure. In the circuit data structure, cell instance structures are formed in addition to interconnection structures and device structures. In the quad tree data structure, a cell layer for bounding boxes of cell instances is formed. The general pre-processing builds a cell structure for each cell where fields for the circuit data structure and the quad tree data structure are formed and the contents of these fields are filled with successive pre-processing when the search of DC path reaches the cell.

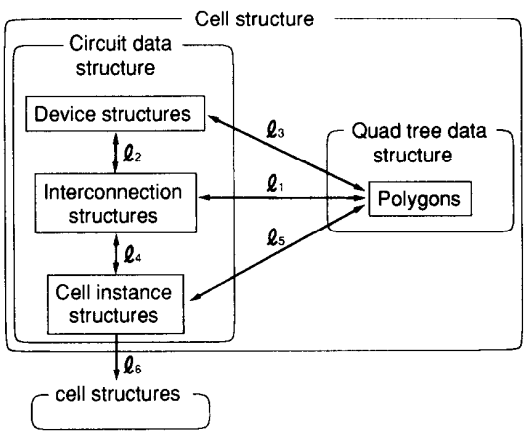


Figure 4. Cell structure used in the hierarchical fault tracing.

4.2.2. Successive pre-processing

The successive pre-processing is executed for a cell layout when the search of DC path reaches the cell. The processing is the same as that in the pre-processing of the previous fault tracing from a flat structured CAD layout data described in subsection 2.2 except that a cell layer for bounding boxes of cell instances in the quad tree data structure and cell instance structures in the circuit data structure are formed and that the interconnections for power supplies are derived by searching interconnections connected to the power supplies in the next higher level cell.

4.2.3. Fault tracing by successive extraction of circuit connectivity information

Here we adapt an extended circuit connection graph where cell instances, MOS transistors and interconnections are vertices and terminals are edges. Each cell has its circuit connection graph.

In order to select upstream interconnections, we uses the *DC PATH LABEL*s put on the DC path where the definition of DC path is extended to include the path which connects a start interconnection to an output terminal of a cell instance. We call this labeling procedure *DC PATH LABEL* labeling. The method allows us to trace a fault hierarchically from the top level cell to the lowest primitive cell and from the primitive cell to the transistor-level circuit in a consistent manner even when the cell data and the transistor-level circuit data exist in a level as a mixture.

Because of the extended definition of DC path, the method is required to judge the terminal of a cell instance is an input terminal or an output one. The input or output judgment can be made by searching a DC path which connects the terminal to the power supplies V_{DD} and V_{SS} in the cell, since it is assumed that the layout does not includes bi-directional interconnections. When the DC path via the examined terminal exists in the cell, the terminal is decided as an output terminal: otherwise it is an input terminal. When there exists cell instances in the cell, the procedure for the search of DC path in the cell is repeated in the next lower level. We call this procedure INPUT/OUTPUT DECISION.

5. APPLICATIONS

We made two programs. One is a program for a circuit logical function extraction which is divided into a transistor-level circuit data extraction program and an automatic circuit data translation program. The other is a program for a hierarchical fault tracing by utilizing a hierarchical layout structure. They are implemented in C programming language on a UNIX workstation (SUN SPARC station 2: 40 MHz, 28.5 MIPS) where the layout data in a GDS-II format is available.

5.1. Circuit logical function extraction

The experimental sample is a CMOS LSI with 300,000 transistors designed with a two metal layer and one polysilicon layer CMOS process. We applied our program to the partial CAD layout data which includes XOR circuits.

The transistor-level netlist extracted by the choice of nine extraction units from the CAD layout data is shown in Fig. 5. The finally obtained netlist and the corresponding circuit schematic are shown in Fig. 6(a) and 6(b), respectively. It is seen that XNOR and XOR gates are recognized correctly.

We measured the CPU times for transistor-level netlist extraction and for circuit data translation as a function of number of extraction units. Results are shown in Fig. 7(a) and 7(b), respectively. From these results, it is seen that either CPU time increases linearly with increasing number of extraction units. The CPU time for the transistor-level netlist extraction is less than one second and the CPU time for the circuit data translation takes several seconds, when the number of extraction units is less than 10: the total processing time amounts to about five seconds which is less than the usual waveform measurement time of several tens of seconds required by the EB test system. Thus it is concluded that the proposed circuit logical function extraction has a sufficient processing time for a hierarchical fault tracing.

5.2. Hierarchical fault tracing by utilizing a hierarchical layout structure

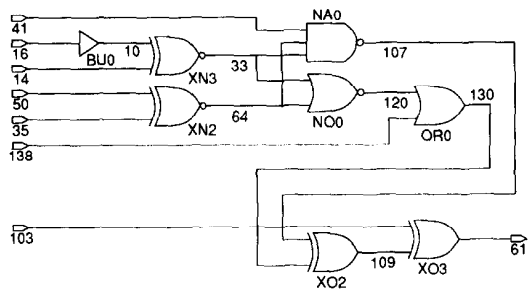
We applied the program to a model layout data of 16-bit full-adder circuit which include about 600 transistors. The layout is designed with a two metal layer and one polysilicon layer

```
*transistor-level netlist.
Vdd 191 0 5V
Vss 536 0 0V
M6762 61 395 536 0 MODEL23
M6938 61 395 191 191 MODEL24
M6936 61 395 191 191 MODEL24
M6757 395 103 783 0 MODEL23
M6759 783 393 536 0 MODEL23
M6753 395 393 790 0 MODEL23
M6751 790 109 536 0 MODEL23
M6934 395 393 191 191 MODEL24
M6931 395 103 1002 191 MODEL24
M6928 1002 109 191 191 MODEL24
M6755 393 103 791 0 MODEL23
M6750 791 109 536 0 MODEL23
M6933 393 103 191 191 MODEL24
M6929 393 109 191 191 MODEL24
M6744 109 425 536 0 MODEL23
M6923 109 425 191 191 MODEL24
M6920 109 425 191 191 MODEL24
M6738 425 107 801 0 MODEL23
M6740 801 424 536 0 MODEL23
M6735 425 424 807 0 MODEL23
M6734 807 130 536 0 MODEL23
M6916 425 424 191 191 MODEL24
M6913 425 107 1018 191 MODEL24
M6910 1018 130 191 191 MODEL24
M6736 424 107 808 0 MODEL23
M6733 808 130 536 0 MODEL23
M6915 424 107 191 191 MODEL24
M6911 424 130 191 191 MODEL24
M6754 107 33 787 0 MODEL23
M6756 787 41 785 0 MODEL23
M6758 785 64 536 0 MODEL23
M6749 107 41 793 0 MODEL23
M6748 793 33 795 0 MODEL23
M6746 795 64 536 0 MODEL23
M6932 107 41 191 191 MODEL24
M6927 107 33 191 191 MODEL24
M6925 107 64 191 191 MODEL24
M6817 33 246 536 0 MODEL23
M6995 33 246 191 191 MODEL24
M6824 246 14 699 0 MODEL23
M6827 699 10 536 0 MODEL23
M6820 246 26 536 0 MODEL23
M7004 246 10 233 191 MODEL24
M6998 233 26 191 191 MODEL24
M7000 246 14 233 191 MODEL24
M6790 64 301 536 0 MODEL23
M6966 64 301 191 191 MODEL24
M6797 301 35 733 0 MODEL23
M6799 733 50 536 0 MODEL23
M6794 301 55 536 0 MODEL23
M6977 301 50 291 191 MODEL24
M6971 291 55 191 191 MODEL24
M6975 301 35 291 191 MODEL24
M6726 130 446 536 0 MODEL23
M6902 130 446 191 191 MODEL24
M6723 446 120 536 0 MODEL23
M6719 446 138 536 0 MODEL23
M6898 446 138 1032 191 MODEL24
M6900 1032 120 191 191 MODEL24
M6741 120 33 536 0 MODEL23
M6737 120 64 536 0 MODEL23
M6917 120 64 1012 191 MODEL24
M6919 1012 33 191 191 MODEL24
M6914 120 33 1016 191 MODEL24
M6912 1016 64 191 191 MODEL24
M6840 10 225 536 0 MODEL23
M6836 10 225 536 0 MODEL23
M7018 10 225 191 191 MODEL24
M7014 10 225 191 191 MODEL24
M6826 26 10 536 0 MODEL23
M6823 26 14 536 0 MODEL23
M7002 26 14 904 191 MODEL24
M7005 904 10 191 191 MODEL24
M6798 55 50 536 0 MODEL23
M6796 55 35 536 0 MODEL23
M6976 55 35 939 191 MODEL24
M6978 939 50 191 191 MODEL24
M6833 225 16 536 0 MODEL23
M7011 225 16 191 191 MODEL24
.MODEL MODEL24 PMOS
.MODEL MODEL23 NMOS
.END
```

Figure 5. The transistor-level netlist extracted by the choice of nine extraction units.

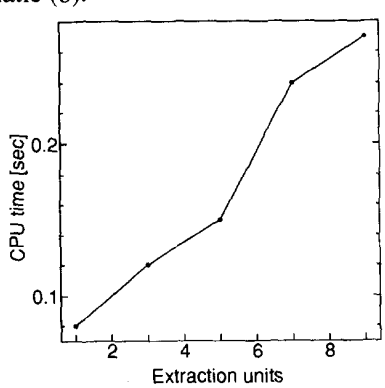
```
*cell-level netlist.
BU0   16   10
NA0    3   33   41   64   107
NO0    2   33   64   120
OR0    2   138  120  130
XN2    2   35   50   64
XN3    2   14   10   33
XO2    2   130  107  109
XO3    2   103  109  61
.END
```

(a)

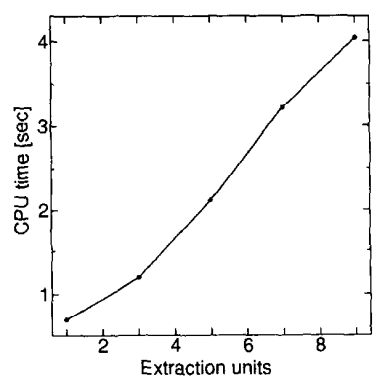


(b)

Figure 6. Finally extracted netlist from the netlist in Fig. 5 (a) and the corresponding circuit schematic (b).



(a)



(b)

Figure 7. CPU times for transistor-level netlist extraction (a) and for circuit data translation (b) measured as a function of number of extraction units.

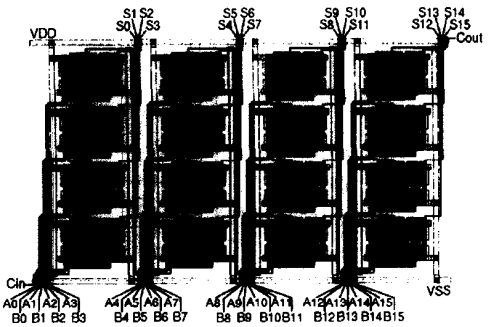


Figure 8. Model layout of a 16-bit full-adder designed with a two metal layer and one polysilicon layer CMOS process.

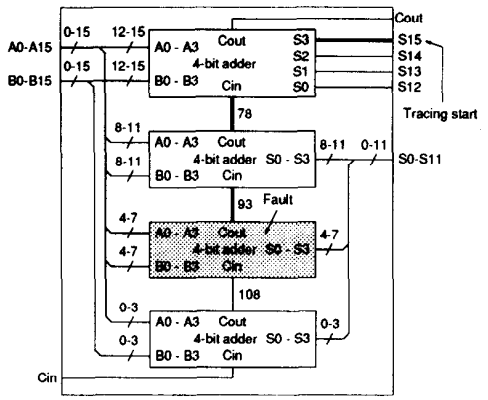


Figure 9. Circuit description which corresponds to the top level cell layout in Fig. 8. The cross hatched cell shows the faulty cell. The thick solid lines indicate a faulty signal propagation path.

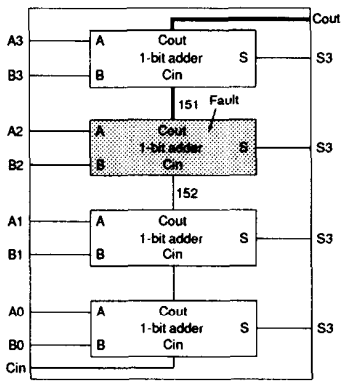


Figure 10. Circuit description which corresponds to a part of the second level cell layout in Fig. 8. The cross hatched cell shows the faulty cell. The thick solid lines indicate a faulty signal propagation path.

CMOS process as shown in Fig. 8. Some components such as wells, substrate contacts, bonding pads are omitted for clarity. In order to simulate the fault tracing, the fault location and the interconnections which propagate a faulty signal were recorded in advance.

Figures 9, 10 and 11 show a 16-bit full-adder circuit, a 4-bit full-adder circuit and a transistor-level circuit for one bit adder, respectively which show a hierarchical structure of the 16-bit adder chip. Though we have only the layout data in CAD database, the circuit diagram is illustrated to explain the situation of a fault propagation. It was assumed that the interconnection 308 in Fig. 11 had a fault and a faulty signal propagated along the interconnections drawn by thick solid lines in Figs. 11, 10 and 9. We traced the fault from the output terminal S15 in Fig. 9. The executed result by the program is shown in Table 1.

The first and the second rows in Table 1 show that faulty signals were detected on the interconnections 78 and 93 which were the input terminals of the first and second 4-bit adders, respectively (see Fig. 9). Signals on other input terminals of these adders are not measured since these terminals are connected to the external input pins of the chip. Next, a good signal was detected on the interconnection 108 (row 3). This indicates that the origin of the fault was in the third 4-bit adder circuit. Then, the fault tracing proceeded to the next lower level, a 4-bit adder, inside the top level cell. After measuring the faulty interconnection 151, a good signals

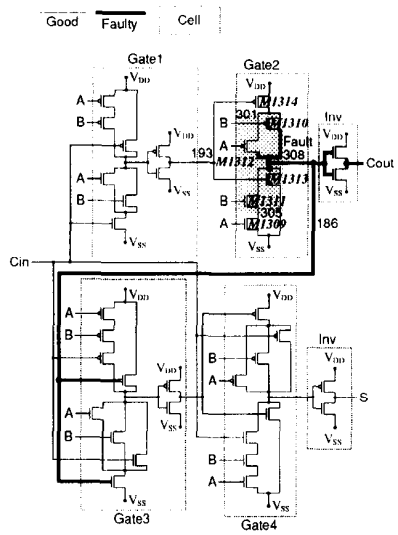


Figure 11. Circuit description which corresponds to a part of the third level cell layout in Fig. 8. The cross hatched region shows the possible fault locations. The thick solid lines indicate a faulty signal propagation path.

Table 1
Executed result

1: measure[1]:78 -> faulty
2: measure[2]:93 -> faulty
3: measure[3]:108 -> good
4: measure[4]:151 -> faulty
5: measure[5]:152 -> good
6: measure[6]:186 -> faulty
7: measure[7]:193 -> good
8: measure[8]:305 -> good
9: measure[9]:301 -> good
10: There may be a fault around the interconnection 308
11: interconnection(308)
12: MOS-FET(1313)
13: MOS-FET(1311)
14: MOS-FET(1312)
15: MOS-FET(1310)

Table 2
Comparison between the efficiency of the hierarchical fault tracing and that of the previous fault tracing from a flat structured layout data.

	CPU time [sec]	Amount of used memory [M bytes]	Number of measured points
Previous fault tracing	58.6	7.28	33
Hierarchical fault tracing	1.52	1.04	9

was measured on the interconnection 152 which was the input terminal of the second 1-bit adder circuit (see Fig. 10). Then the fault tracing proceeded to the next lower level, a 1-bit adder, inside that cell. After measuring the faulty interconnection 186, a good signal was measured on the interconnection 193 which was the input to the Gate 2 (see Fig. 11). Then the fault tracing proceeded to the lowest level or a transistor-level circuit. Since good signals were detected on the interconnections 305 and 301 which were upstream interconnections of the interconnection 308 or the interconnection 186, it was concluded that there may be a fault around the interconnection 308 (the interconnection 186). Rows from 11 to 15 suggest the names of possible faulty interconnection and devices. These results are shown in the crosshatched region in Fig. 11. Thus, the intended fault was specified.

To compare the efficiency of the hierarchical fault tracing with that of the previous fault tracing from a flat structured layout data, we applied the previous non-hierarchical method to the flattened layout data of 16-bit adder circuit. The result is shown in Table 2 where the CPU time,

the amount of used memory and the number of measured points are compared. The CPU time does not include EB probing time. From this table, the hierarchical fault tracing improves the efficiency of the previous method by a factor of about 38 in CPU time, 7 in the amount of used memory and 3.7 in the number of measured points.

6. CONCLUSIONS

We have proposed two approaches for efficient fault diagnosis where only CAD layout data is available in the CAD-linked electron beam test system.

One is a circuit logical function extraction method from the CAD layout data. The method consists of two steps: 1) transistor-level circuit data extraction from the CAD layout data by utilizing the successive circuit extraction algorithm, and 2) automatic circuit data translation of transistor-level data to higher cell-level data hierarchically. In the transistor-level circuit data extraction, the circuit data of a primitive logic gate is taken as an extraction unit for convenience in a hierarchical fault tracing. A part of the circuit data is extracted by the choice of the number of extraction units along DC paths in the CAD layout on demand of the hierarchical fault tracing algorithm. The circuit data translation adapts a hybrid approach that combines the logical synthesis technique and the library-based graph-isomorphism technique. The method gives us the netlist where the net name is unique in spite of the level in hierarchy. We applied the method to the CAD layout data of a CMOS LSI. The results show that the proposed circuit logical function extraction method has a sufficient processing time for a hierarchical fault tracing.

The other is a hierarchical fault tracing method by utilizing a hierarchical layout structure. The method allows us to trace a fault hierarchically from the top level cell to the lowest primitive cell and from the primitive cell to the transistor-level circuit in a consistent manner independent of circuit functions even when the cell data and the transistor-level circuit data exist in a level as a mixture. The method is applied to a hierarchically structured CMOS model layout with about 600 transistors. The results show that the hierarchical fault tracing improves the efficiency of the previous non-hierarchical method by a factor of about 38 in CPU time, 7 in the amount of used memory and 3.7 in the number of measured points.

REFERENCES

1. Director, S. W., Maly, W. and Strojwas, A. J., VLSI Design for Manufacturing : Yield Enhancement, Kluwer Academic Publishers, Boston, 1990.
2. Ura, K. and Fujioka, H., "Electron beam testing," Advances in Electronics and Electron Physics, ed. P. W. Hawkes, Academic Press, New York, Vol. 73, pp. 223–317, 1989.
3. Noble, A. C., "IDA: A tool for Computer-Aided Failure Analysis," Proc. 1992 international Test Conference, pp. 848–853, 1992
4. Yamaguchi, N., Sakamoto, T., Nishioka, H., Majima, T., Satou, T., Shinada, H., Todokoro, H. and Yamada, O., "E-Beam Fault Diagnosis System for Logic VLSIs," Microelectronic Engineering, 16, pp. 121–128, 1992.
5. Kuji, N. and Shirakawa, K., "Cone/Block Methods for Logic Simulation Time Reduction in E-Beam Guided-Probe Diagnosis," IEICE Trans. Electron., **E77–C**, pp. 560–566, 1994.
6. Miura, K., Nakamae, K. and Fujioka, H, "Automatic Tracing of Transistor-Level Performance Faults with CAD-Linked Electron Beam Test System," IEICE Trans. Fundamentals, **E77–A**, pp. 539–545, 1994.

7. Fujioka, H. and Nakamae, K., "LSI Failure Analysis with CAD-Linked Electron Beam Test System and Its Cost Evaluation," IEICE Trans. Electronics, **E77–C**, pp. 535–545, April 1994.
8. Norimatsu, K., Shido, M., Ishikawa, M. and Fujii, M., "Development of the Navigation System for Gate Level Design and Its Application to VLSIs," in Proc. of Symposium on Electron Beam Testing, Osaka: Japan Society for the Promotion of Science, pp. 61–70, 1992.
9. Miura, K., Nakamae, K. and Fujioka, H., "Automatic Transistor-Level Performance Fault Tracing by Successive Circuit Extraction from CAD Layout Data for VLSI in the CAD-Linked EB Test System," IEICE Trans. Electronics, **E78–C**, 1995, to be published.
10. Kishimoto, A., Kawanishi, H., Yoshizawa, H., Ohno, H., Fujinami, Y. and Kani, K., "An Interconnection Check Algorithm for Mask Pattern," in Proc. IEEE Conf. of ISCAS, pp. 669–672, 1979.
11. Watanabe, T., Endo, M. and Miyahara, N., "A New Automatic Logic Interconnection Verification System for VLSI Design," IEEE Trans. Computer-Aided Design, **CAD–2**, pp. 70–81, 1983.
12. Kostelijk, T. and Loore, B. D., "Automatic Verification of Library-Based IC Designs," IEEE Journal of Solid-State Circuits, **26**, pp. 394–403, 1991.
13. Trimberger, S. M., An Introduction to CAD for VLSI, Kluwer Academic Publishers, Boston, 1987.
14. West, N. H. E. and Eshraghian, K., Principles of CMOS VLSI Design: Second Edition, Addison-Wesley Publishing, New York, 1993.