1993

The Design of High-Frequency Continuous-Time Integrated Analog Signal Processing Circuits

Pan Wu
Portland State University

Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds
Let us know how access to this document benefits you.

Recommended Citation

10.15760/etd.1161

This Dissertation is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.
THE DESIGN OF HIGH-FREQUENCY CONTINUOUS-TIME INTEGRATED
ANALOG SIGNAL PROCESSING CIRCUITS

by

PAN WU

A dissertation submitted in partial fulfillment of
the requirements for the degree of

DOCTOR OF PHILOSOPHY
in
ELECTRICAL AND COMPUTER ENGINEERING

Portland State University
1993
DISSERTATION APPROVAL

The abstract and dissertation of Pan Wu for the Doctor of Philosophy in Electrical and Computer Engineering were presented on Dec. 10, 1993, and accepted by the dissertation committee and the doctor program.

COMMITTEE APPROVALS:

Rolf Schaumann, Chair

Marek Berkowski

Malgorzata Chrzanowska-Jeske

Stewart Taylor

Pavel Smejtek
Representative of the Office of Graduate Studies

DOCTORAL PROGRAM APPROVAL:

Rolf Schaumann, Chair
Department of Electrical Engineering

******************************************************************************************

ACCEPTED FOR PORTLAND STATE UNIVERSITY BY THE LIBRARY

by on March 3, 1994
ABSTRACT


Title: The Design of High-Frequency Continuous-Time Integrated Analog Signal Processing Circuits.

High-performance, high-frequency operational transconductance amplifiers (OTAs) are very important elements in the design of high-frequency continuous-time integrated analog signal processing circuits, because resistors, inductors, integrators, mutators, buffers, multipliers, and filters can be built by OTAs and capacitors. The critical considerations for OTA design are linearity, tuning, frequency response, output impedance, power supply rejection (PSR) and common-mode rejection (CMR).

For linearity considerations, two different methods are proposed. One uses cross-coupled pairs (CMOS or NMOS), producing OTAs with very high linearity but either the input range is relatively small or the CMR to asymmetrical inputs is poor. Another employs multiple differential pairs (current addition or subtraction), producing OTAs with high linearity over a very large input range. So, there are tradeoffs among the critical considerations. For different applications, different OTAs should be selected.

For consideration of frequency response, the first reported GaAs OTA was designed for achieving very-high-frequency performance, instead of using AC compensation techniques. GaAs is one of the fastest available technologies, but it was new and
less mature than silicon when we started the design in 1989. So, there were several issues, such as low output impedance, no P-channel devices, and Schottky clamp. To overcome these problems, new techniques are proposed, and the designed OTA has comparable performance to a CMOS OTA.

For PSR and CMR considerations, a fully balanced circuit structure is employed with a common-mode feedback (CMF) circuit used to stabilize the DC output voltages. To reduce the interaction of the operation of CMF and tuning of OTAs, three improved versions of the CMF circuits used in operational amplifiers are proposed.

With the designed OTAs, a 1 GHz GaAs inductor with small parasitics is designed using the proposed procedure to reduce high-frequency effects. Two CMOS high-order, high-frequency filters are designed: one in cascade structure and one in LC ladder form. Also, a 200 MHz third-order elliptic GaAs filter is designed with special consideration of very-high-frequency parasitics.

All circuits were fabricated and measured. The experimental results were used to verify the designs.
ACKNOWLEDGEMENT

I consider myself very fortunate to be the first Ph. D. student of Prof. Rolf Schuermann at Portland State University. He has near twenty years' teaching and research experience at the University of Minnesota and is a leader in the field of the design of analog integrated signal processing circuits. With his guidance, advice, suggestions, countless discussions, and encouragement, the success of this dissertation became true. Also, affected by his contributions and academic spirit, I came to work harder and harder, and had the opportunity to publish 16 papers in international journals and conferences for this research and obtained 4 awards.

I owe special thanks to Prof. Marek Perkowski who is the first person teaching me how to do research and encouraging me to work hard. Special thanks also to Profs. Malgorzata Chrzanowska-Jeske and Pavel Smejtek for their important and useful suggestions to my Ph.D program.

I also feel lucky to have Dr. Stewart Taylor as my Ph. D. dissertation committee member. He has over ten years' design experience at Tektronix Inc. and TriQuint Semiconductor, Inc. He let me have a chance to learn what is interesting to industry and how to combine theory and practice. He also taught me many design skills with GaAs technology, and laid out and fabricated my GaAs OTA in TriQuint's fabrication process without charge.

Besides my Ph. D. dissertation committee members, I am also truly appreciative of Mr. Paul Latham, Design Section Head in Allegro Microsystems, Inc., for his numerous suggestions to one of my CMOS OTAs and filter designs and their free fabrication. Grateful appreciation is also due to Prof. Robert Daasch for his help with layout.
design and fabrication of the circuits in the MOSIS process. For chip measurement, I would like to thank Mr. Tom Kautzman and Profs. Richard Tymerski and Paul Van Halen for their support with the equipment.

For financial support that allowed me to complete this dissertation, I would like to acknowledge Tektronix, Inc., for the "Tektronix Fellowship", Portland State University for a Research Assistantship from "the Faculty Development Fund" and a Teaching Assistantship, and Top-Vu Technology, Inc., for a consultant job.

A very special thanks to my wife, Lili Yan, for her long-time support, without which the completion of this dissertation would be impossible. Her help on the figures for this dissertation and other papers is greatly appreciated. Also, special thanks and respect to Ms. Shirley Clark for her administrative support and special help.

Finally, I owe thanks to Mr. Joseph Schutz and Mr. Sudarshan Kumar from Intel Corp. for their understanding and support to finish this dissertation.
TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>CONTENT</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II</td>
<td>HIGH-LINEARITY CMOS OTAs</td>
<td>14</td>
</tr>
<tr>
<td>II.1</td>
<td>Basic Principles</td>
<td>14</td>
</tr>
<tr>
<td>II.1.1</td>
<td>Transconductor as a Basic Building Block</td>
<td>14</td>
</tr>
<tr>
<td>II.1.2</td>
<td>Transconductor Based on a Single MOSFET</td>
<td>16</td>
</tr>
<tr>
<td>II.1.3</td>
<td>Transconductor Based on Differential Pair</td>
<td>19</td>
</tr>
<tr>
<td>II.1.4</td>
<td>General Overview</td>
<td>21</td>
</tr>
<tr>
<td>II.2</td>
<td>Transconductor Based on Cross-Coupled Pairs</td>
<td>23</td>
</tr>
<tr>
<td>II.2.1</td>
<td>CMOS Cross-Coupled Pairs</td>
<td>23</td>
</tr>
<tr>
<td>II.2.2</td>
<td>NMOS Cross-Coupled Pairs</td>
<td>32</td>
</tr>
<tr>
<td>II.2.3</td>
<td>Conclusions</td>
<td>36</td>
</tr>
<tr>
<td>II.3</td>
<td>Transconductor Based on Multiple Differential Pairs</td>
<td>44</td>
</tr>
<tr>
<td>II.3.1</td>
<td>Basic Principle</td>
<td>45</td>
</tr>
<tr>
<td>II.3.2</td>
<td>Practical Design</td>
<td>49</td>
</tr>
<tr>
<td>II.3.3</td>
<td>Other Considerations</td>
<td>57</td>
</tr>
<tr>
<td>II.4</td>
<td>Comparisons</td>
<td>64</td>
</tr>
<tr>
<td>Section</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>III</td>
<td>DESIGN CONSIDERATIONS FOR COMMON-MODE FEEDBACK</td>
<td>66</td>
</tr>
<tr>
<td></td>
<td>III.1 Basic Principles</td>
<td>67</td>
</tr>
<tr>
<td></td>
<td>III.2 Improved CMF Circuits</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>III.3 Discussion</td>
<td>80</td>
</tr>
<tr>
<td>IV</td>
<td>CMOS CONTINUOUS-TIME INTEGRATED FILTERS</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>IV.1 The Basic Principles</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>IV.1.1 Cascade Filter Design</td>
<td>84</td>
</tr>
<tr>
<td></td>
<td>IV.1.2 LC Ladder Filter Design</td>
<td>91</td>
</tr>
<tr>
<td></td>
<td>IV.2 A 20 MHz Third-Order Elliptic CMOS Filter</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>IV.2.1 Filter Design</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>IV.2.2 Design Considerations for Frequency Responses</td>
<td>98</td>
</tr>
<tr>
<td></td>
<td>IV.2.3 Simulation Results</td>
<td>102</td>
</tr>
<tr>
<td></td>
<td>IV.3 A 4 MHz Sixth-Order Bessel CMOS Filter</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>IV.3.1 Filter Design</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>IV.3.2 Design Considerations and Simulation Results</td>
<td>112</td>
</tr>
<tr>
<td>V</td>
<td>ASP CIRCUITS WITH VERY-HIGH-FREQUENCY RESPONSE</td>
<td>118</td>
</tr>
<tr>
<td></td>
<td>V.1 A 7 GHz GaAs OTA</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>V.1.1 Linearity</td>
<td>120</td>
</tr>
<tr>
<td></td>
<td>V.1.2 Output Impedance</td>
<td>125</td>
</tr>
<tr>
<td></td>
<td>V.1.3 Tuning Capability</td>
<td>128</td>
</tr>
<tr>
<td></td>
<td>V.1.4 DC level shifter</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>V.1.5 Common-Mode Feedback Circuit</td>
<td>136</td>
</tr>
<tr>
<td></td>
<td>V.1.6 Simulation Results</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td>V.1.7 Other Considerations</td>
<td>139</td>
</tr>
<tr>
<td></td>
<td>V.2 A 1 GHz GaAs Gyrator-C Inductor</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>V.2.1 Design Considerations for Gyrator-C Inductors</td>
<td>143</td>
</tr>
</tbody>
</table>
V.2.2 Practical Design .............................................. 148

V.3 A Simulated 200 MHz GaAs LC Ladder Filter ................. 150

VI IC PROCESSES AND LAYOUT DESIGNS ...................... 153

VI.1 Description of IC Processes .................................. 154

VI.1.1 MOSIS ....................................................... 154
VI.1.2 DABIC ....................................................... 155
VI.1.3 QED ......................................................... 158

VI.2 Description of the Layout Tool ................................ 158

VI.3 Design Considerations for Layout .............................. 160

VI.3.1 Parasitics ................................................... 160
VI.3.2 Process Tolerances ......................................... 166
VI.3.3 Reliability ................................................... 171
VI.3.4 Design Automation .......................................... 172

VI.4 Layout Designs ................................................. 173

VI.4.1 Layout in MOSIS ............................................ 173
VI.4.2 Layout in DABIC ............................................ 181
VI.4.3 Layout in QED .............................................. 186

VI.5 Post-Layout Performance Verifications ....................... 189

VII EXPERIMENTAL RESULTS .................................... 192

VII.1 Chip from MOSIS ............................................. 192

VII.1.1 CMOS OTA with Current Addition ....................... 192
VII.1.2 CMOS Elliptic Filter ..................................... 195

VII.2 Chip from Allegro Microsystems, Inc. ......................... 201

VII.2.1 CMOS OTA with Current Subtraction .................... 201
VII.2.2 CMOS Bessel Filter ...................................... 202

VII.3 GaAs OTA from TriQuint Semiconductor, Inc. ............... 208
## LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Summary of the Comparison of the Different OTAs Designed in This Chapter</td>
</tr>
<tr>
<td>II</td>
<td>Summary of the Comparison of Three Typical CMF Circuits</td>
</tr>
<tr>
<td>III</td>
<td>Summary of the Comparison of the Different CMF Circuits</td>
</tr>
<tr>
<td>IV</td>
<td>Typical Performance Characteristics of the CMOS Elliptic Filter</td>
</tr>
<tr>
<td>V</td>
<td>Experimental Data for the Output Voltages to the Load Resistors</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Transconductor as a Floating Voltage-Controlled Current Source.</td>
</tr>
<tr>
<td>2.</td>
<td>Resistors Built by Transconductor.</td>
</tr>
<tr>
<td>3.</td>
<td>A Voltage Amplifier (a) and an Integrator (b) Built by Transconductor.</td>
</tr>
<tr>
<td>4.</td>
<td>A Gyrator (a) and an Inductor (b) Built by Transconductors.</td>
</tr>
<tr>
<td>5.</td>
<td>A Symbol (a) and a Small-Signal AC Model (b) of an MOSFET.</td>
</tr>
<tr>
<td>6.</td>
<td>A General Unsymmetrical Differential Pair.</td>
</tr>
<tr>
<td>7.</td>
<td>Linear Transconductor Circuit Base on CMOS Cross-Coupled Pairs.</td>
</tr>
<tr>
<td>8.</td>
<td>Modified Arrangement of the Transconductor Circuit in Fig. 7.</td>
</tr>
<tr>
<td>9.</td>
<td>Complete Circuit Diagram of the Circuit in Fig. 8.</td>
</tr>
<tr>
<td>10.</td>
<td>SPICE Simulation Results for the Circuit in Fig. 9.</td>
</tr>
<tr>
<td>11.</td>
<td>Linear Transconductor Circuit Base on NMOS Cross-Coupled Pairs.</td>
</tr>
<tr>
<td>12.</td>
<td>The Simulation Results for the OTA in Fig. 11. (a) DC Transfer Curves and (b) Tunability.</td>
</tr>
<tr>
<td>13.</td>
<td>The Simulation Results for the Frequency Responses of the OTA in Fig. 11.</td>
</tr>
<tr>
<td>14.</td>
<td>SPICE Simulation Results for the OTA in Fig. 11 with an Asymmetrical Input Signals: $V_1 = V_{in}$ and $V_2 = 0$.</td>
</tr>
<tr>
<td>15.</td>
<td>The Simulation Results for the Total Output Current and Transconductance of the OTA in Fig. 7 with $V_1 = V_{in}$ and $V_2 = 0$.</td>
</tr>
<tr>
<td>16.</td>
<td>SPICE Simulation Results for the OTA in Fig. 11 with an Asymmetrical Input Signals: $V_1 = 1.5V_{in}$ and $V_2 = 0.5V_{in}$.</td>
</tr>
</tbody>
</table>
17. The Simulation Results for the Total Output Current and Transconductance of the OTA in Fig. 7 with \( V_1 = 1.5V_{in} \) and \( V_2 = 0.5V_{in} \). ........................................ 44

18. SPICE Simulation Results for the Output Current of a General Unsymmetrical Differential Pair in Fig. 6. ................................................................. 48

19. Linear Transconductance Circuit Based on Multiple Differential Pairs. .. 50

20. SPICE Simulation Results for the Transconductances \( g_{m1}, g_{m2}, g_{m3} \), and \( g_{m1} + g_{m2} \) in the OTA in Fig. 19. .................................................... 52

21. The Simulation Results for the Circuit in Fig. 19. (a) The Transfer Curve and (b) Linearity Error. ................................................................. 54

22. SPICE Simulation Results for the Transconductances \( g_{m1}, g_{m2}, g_{m3} \), and \( g_{m1} + g_{m2} \) in the OTA with Current Subtraction. ........................................ 56

23. The Simulation Results for the Transfer Curve and the Transconductance of the OTA with Current Subtraction. .................................................... 58

24. SPICE Simulation Results for OTA Tuning of the Magnitude of \( g_m \) by Varying \( I_s \) in Fig. 19. (a) Current Addition and (b) Current Subtraction. 59

25. A RC Branch Used for OTA Tuning of the Phase of \( g_m \); (a) with Passive Resistor, and (b) with Transistor. (c) SPICE Simulation Results. .......... 61

26. SPICE Simulation Results for Frequency Responses of the OTA in Fig. 19 with Tuning. ................................................................................. 62

27. The Simulation Results for Output Impedance of the OTA in Fig. 19 with Tuning. ................................................................................. 63

28. The Three Most Common CMF Circuits Used in Op-Amps. ............... 69

29. SPICE Simulation Results for Common-Mode Output Voltages, \( V_{o+} \) and \( V_{o-} \) with Different \( V_{so} \). ......................................................... 72

30. Circuit Diagram for a Simple Fully Differential OTA with a Typical Common-Mode Feedback Circuit. .......................................................... 74
| 31. | Circuit Diagram for the Improved CMF Circuits Suitable for OTAs with Tuning; First Choice. | 76 |
| 32. | SPICE Simulation Results for Common-Mode Output Voltage, $V_{o+}$, of the OTA with the CMF Circuits <1> in Fig. 30 and <2> in Fig. 31. | 77 |
| 33. | Circuit Diagram for the Improved CMF Circuits Suitable for OTAs with Tuning; Second Choice. | 79 |
| 34. | Circuit Diagram for the Improved CMF Circuits Suitable for OTAs with Tuning; Third Choice. | 79 |
| 35. | The Block Diagram of a Filtering Network (a) and Cascade Connection (b). | 85 |
| 36. | The Filter Characteristics of Lowpass (a), Highpass (b), and Bandpass (c). | 86 |
| 37. | The Biquadratic Filters of Two-Integrator Loop (a) and Parallel RLC (b). | 88 |
| 38. | Fully Differential OTA-C Implementations of the Biquadratic Filters in Fig. 37a (a) and in Fig. 37b (b). | 89 |
| 39. | Resistively Terminated Lossless Two-Port Network. | 92 |
| 40. | Resistively Terminated Lossless LC Ladder Network. | 94 |
| 41. | Circuit Diagram of a LC Ladder Filter. | 94 |
| 42. | Fully Differential OTA-C Implementations of the Basic Building Blocks in ASP Circuits. A Voltage to Current Converter (a), a Resistor (b), an Unity-Gain Buffer (c), and a Gyrator-C Inductor (d). | 95 |
| 43. | A Third-Order Elliptic Low-Pass Filter with Passive Elements. | 97 |
| 44. | Circuit Diagram of the Filter in Fig. 43 with OTAs and Capacitors. | 99 |
| 45. | Circuit Diagram of CMOS Linear OTA with Cascode Output Stage. | 103 |
| 46. | Input Impedance (a) and Output Impedance (b) of the OTA in Fig. 45. | 105 |
47. Total Output Current and Transconductance of the OTA in Fig. 45. 106
48. The Transfer Function of the Third-Order Lowpass Elliptic Filter with Tuning (a) and Its Details (b). 106
49. Circuit Structure of Sixth-Order Cascade Fully Differential OTA-C Bessel Filter. 110
50. Output Stage of the Filter in Fig. 49. 111
51. SPICE Simulations of the Gain <1>, Delay <2> and Phase <3> Responses of the Filter in Fig. 49. 117
52. A Matched Differential-Pair with Source Degeneration. 121
53. SPICE Simulation Results for the Output Current as a Function of the Resistor $R_E$ in the Circuit in Fig. 52. 123
54. GaAs Amplifiers with Cascode Driver and "Self-Bootstrapping" Load (a) and the "Improved Self-Bootstrapping" Load (b). 127
55. Transconductance Tuning by Using Shunt Current Source (a) or Transistor $Q_b$ (b). 130
56. SPICE Simulation Results for DC Transfer Curves with Tuning:

\[-0.8 \, V \leq V_a \leq 0.3 \, V.\] 132
57. Transconductance Tuning by Using the Combination of "Digital" and "Analog" Techniques. 133
58. The Output Buffer or Level Shifter. 136
59. Circuit diagram for a GaAs OTA. 137
60. SPICE Simulation Results for Frequency Response of the GaAs OTA with Different Tuning Voltages (a) and Output Impedances (b). 140
61. The Simulation Results for AC Compensation of the Excess Phase Shifts by $C_1$ and $C_2$ in the Circuit in Fig. 59. 141
62. The Simulation Results for the Output Current of the OTA in Fig. 59:
<1> T = 27°C without wire inductors, <2> T = 100°C without wire inductors, and <3> T = 27°C with 2 nH wire inductors.

63. Small-Signal Equivalent Circuit for Single-Ended OTA. .................................................. 143

64. A Equivalent Circuit for the Gyrator-C Inductor Shown in Fig. 42d.
(a) RLC Equivalent Network. (b) Small-Signal Circuit for Single-Ended Structure. .......................... 144

65. A Simplified Circuit for the First OTA Used in Gyrator-C Inductor. ...... 147

66. SPICE Simulated Frequency Response of a Gyrator-C Inductor. .................. 149

67. The Simulation Results of the Frequency Responses with Tuning for the Filter in Fig. 44, but with GaAs OTAs in Fig. 59. .................................................. 152

68. SPICE Simulation Results for the Drain Current vs. Drain-To-Source Voltage Characteristics of an N-MOSFET (a) and a P-MOSFET (b) from MOSIS. .................................................. 156

69. The Simulation Results for the Drain Current vs. Drain-To-Source Voltage Characteristics of an N-MOSFET (a) and a P-MOSFET (b) from Allegro Microsystems, Inc. .................................................. 157

70. SPICE Simulation Results for the Drain Current vs. Gate-To-Source Voltage Characteristic (a) and the Drain Current vs. Drain-To-Source Voltage Characteristic (b) of a GaAs NMESFET from TriQuint Semiconductor, Inc. .................................................. 159

71. A Cross-Section of the Layout of a p-well CMOS Structure (a) and Its Top View (b). .................................................................................................................. 162

72. Basic Latch Loop Formed by Parasitic Bipolar Transistors and Parasitic Resistors in CMOS Structure. ........................................................................ 162

73. Layout of Two MOS Transistors with the Shared Diffusion (Shaded Area). .............................................................................................................. 164
74. Layouts of a MOS Transistor with One Leg (a) and Three Legs (b). 165
75. Design Considerations of the Typical Integrated Capacitors for the Edge
   Effects (a) and Capacitance Ratio-Matching (b). 170
76. Layer Names and Colors of MOSIS Process. 174
77. Layout of the CMOS OTA with Current Addition Shown in Fig. 45. 176
78. Microphotograph of the Third-Order Elliptic Filter (see Fig. 44)
   Fabricated by the MOSIS Vendor. 179
79. The Equivalent CMOS ESD Pin Protection Circuit Built by P+ and N+
   Diffusions (a) and the Cross-Section of the Layout (b). 181
80. Layout of the CMOS OTA with the Current Subtraction Method Shown
   in Fig. 6.14. 183
81. Circuit Diagram for the OTA (Current Subtraction) with Bias Circuit to
   Common-Mode Feedback Circuit. 184
82. Microphotograph of the Sixth-Order Bessel Filter (see Fig. 49)
   Fabricated by Allegro Microsystems, Inc. 185
83. ESD Pin Protection Circuit Used in DABIC Chip Fabricated by Allegro
   Microsystems, Inc. 187
84. Layout of ESD Pin Protection Circuit Shown in Fig. 83. 187
85. Layout of the GaAs OTA Shown in Fig. 59. 188
86. Microphotograph of the GaAs OTA Fabricated by TriQuint
   Semiconductor, Inc. 190
87. SPICE Simulation Results Based on the OTA Circuit (Fig. 45)
   Extracted from the Layout Shown in Fig. 77. 191
88. Experimental Setup for the DC Measurements. 193
89. Experimental Results for the DC Transfer Curve of the MOSIS OTA
   with Current Addition. 194
90. Experimental Setup for the Measurement of the High-Frequency Filter. 196
91. Experimental Results for the Frequency Response of the Third-Order Elliptic Filter. 197
92. Measured Spectrum for the Third-Order Elliptic Filter; Input (a) and Output (b). 198
93. Total Harmonic Distortion vs the Differential Input Peak-To-Peak Voltage. 199
94. Experimental Results for the Frequency Responses of the Third-Order Elliptic Filter with Tuning. 200
95. Experimental Results for the DC Transfer Curve of the OTA with Current Subtraction. 202
96. Total Harmonic Distortion vs the Differential Input Peak-To-Peak Voltage for the OTA with Current Subtraction. 203
97. Experimental Results of the AC Transfer Function and Noise Performance for the Sixth-Order Bessel Filter. 206
98. Experimental Results of the Phase as Function of Frequency for the Sixth-Order Bessel Filter. 207
99. Phase of the Transformer as Function of Frequency. 207
100. Experimental Results for the Frequency Responses of the Bessel Filter with Tuning. 208
101. Experimental Results for the DC Transfer Curve of the GaAs OTA. 209
102. Small-signal model of an OTA with the load resistor $R_L$. 210
103. Experimental Results for the Frequency Response of the GaAs OTA. 212
CHAPTER I

INTRODUCTION

In recent years, with the growing industrial demand for signal processing systems operating at higher and higher frequencies, more and more attention is being paid to analog signal processing (ASP) techniques and circuits, particularly, to circuits operating in the continuous-time domain. It stands to reason that high-frequency signal processing problems are often solved more effectively in a continuous-time analog environment than in the sampled-data analog (switched-capacitor) or digital domains, because there is no need for switching circuitry, sampling, or A/D and D/A conversion. The other advantages of continuous-time analog signal processing are lower power consumption, no switching noise and no aliasing problems [1-11]. Consequently, considerable research has been performed on the design of ASP circuits and these circuits find more and more applications in many fields, such as filters, neural networks, communication systems, computer disk drives and read-out and preprocessing electronics in detector arrays.

In analog signal processing systems, to select a signal within a required frequency range is a very important task which is, generally, implemented using electronic filters [1, 2, 12, 13]. For instance, in radios, TVs, telephony, communication systems, instrumentation systems, and personal computers, electronic filters are basic and unavoidable components.

From 1920 to the late 1960s, electronic filters were mainly designed as resistively terminated lossless LC filters using passive components, such as resistors, capacitors, and inductors [1, 2, 12, 14-16]. After the middle 1960s, with high-quality operational amplifiers (op-amps) becoming economically available [17-19], active RC filters started
to replace the passive filters in order to avoid inductors because the latter usually are heavy and expensive, have a large volume, and are hard to integrate [20-28]. In the late 1970s, miniaturization and integration continued and resulted in "resistorless" active filters: switched-capacitor filters [29-41]. Since the middle 1980s, with success in the design of linear transconductance amplifiers (labeled as transconductor) [4, 42-70] and linear transresitors [71-84], transconductance-C filters (where C stands for capacitance) and MOSFET-C filters have been available.

Among the above active filters, RC-active filters become less and less important because of poor bandwidth and large chip die area for passive resistors in fully integrated circuits [85-91]. In switched-capacitor filters, a resistor is replaced by one capacitor and two or more switches which are implemented by transistors with a sampling clock, so chip area is small. The additional achievements for this replacement are small process tolerances and low power dissipation [33-35, 92-96]. Therefore, the design of switched-capacitor filters was a major improvement in the implementation of analog integrated filters during last ten years. Unfortunately, the operating frequencies of these filters are limited by sampling frequency and by relatively poor frequency responses of op-amps [1, 4, 5, 97-101]; these filters, therefore, generally have good performances only in the kilohertz frequency range which is not high enough to meet frequency requirements of high-speed systems. Furthermore, switched-capacitor filters generally need a couple of continuous-time filters (antialiasing and reconstruction) as interfaces to the real world [1, 39, 102]; these are often the weakest link in the filter chain.

Among different classes of filters, continuous-time filters are the only candidates able to use the whole frequency range that an available technology can offer [1-3, 103-105]. MOSFET-C filters are continuous-time filters in which the resistors are replaced by transistors [106-112]. But, because they are also op-amp based filters, their frequency responses are still restricted by the poor bandwidth of the op-amps used [1, 98-99, 101,
Transconductance-C filters are another kind of continuous-time filter in which there are neither sampling clocks nor op-amps [1, 3, 6, 119-120]. The frequency performance of this kind of filter, therefore, is only limited by the frequency response of transconductors used, which, generally, is much better than that of op-amps and can be designed to be as high as two-thirds of the transition frequency of the transistors [97, 121-125]. In recent years, the design of transconductance-C filters became the dominant technique for high-frequency filter design (above 5 MHz) [102, 126-137] and commercial circuits are available [120, 135, 136].

The method to build transconductance-C filters is to use one transconductor to simulate a passive resistor, and use two transconductors and one capacitor to simulate a passive inductor [1, 21, 138-142]. Alternatively, integrators, consisting of one transconductor and one capacitor, are used to simulate the action of inductors and capacitors [1, 48, 119, 143-145]. In both cases, this kind of filter contains only transconductors and capacitors. This results in the overall performance of the filter being critically determined by that of the transconductors used, and the design of a high-performance transconductor becomes an important task.

The history of development of electronic filters shows that to design a high-performance transconductor is very critical for high-speed filter design. Besides filters, many other ASP circuits can be built by transconductors and capacitors. Among the examples are signal amplifier, output buffer, integrator, gyrator, analog multiplier, negative impedance, etc. [1, 48, 138, 139, 142, 147-152]. Therefore, this dissertation will focus mainly on the design of high-performance transconductors, and high-frequency, high-order integrated filters will be designed as examples of ASP circuit design using transconductors and capacitors. Of course, design methods for improving the high-frequency performance by reducing the high-frequency parasitics will be proposed for high-frequency filter design.
The purpose of designing high-frequency filters in this dissertation is not only to show an application of the designed high-performance transconductors, but also to set up a measurement environment for transconductors. In practice, to measure the frequency response of transconductors operating above hundred megahertz is very difficult, because the parasitic pole formed by pin capacitances and high output impedance of the transconductors will become the dominant pole if pin capacitance is larger than few tenths of pico-farad; this usually happens in present technologies [134, 153]. If a small load is used to reduce the impact of the high output impedance, it results in a mismatch between the transconductor and the load. In other words, the output parameter of a transconductor is current which, usually, cannot be directly measured with the present equipment that is mostly designed to measure voltage instead of current. To use an output buffer converting the output current to voltage does generally not help because the relationship between the output voltage and the current is very hard to define. Also, the voltage gain usually is small if the load or the input impedance of the equipment is small, as is generally true for high-frequency equipment. For example, the input impedance of the HP gain-phase analyzer (HP4194A) is designed as 50 Ω to match transmission lines. Then, signal-to-noise ratio and spectral content are hard to measure. For filters there are no the above-mentioned problems for measurements because the measured parameter is the output voltage with, generally, unity gain to a 50 Ω load [1, 7, 102, 127, 128, 135]. On the other hand, as mentioned above, the performance of transconductance-C filters is determined by the transconductors used; a high-quality filter will prove the performance of the transconductors.

Another advantage of transconductance-C based ASP circuits is that these circuits can be designed by computer-aided design methods. In recent years, with rapid development of integrated circuits, design automation has led to the design of very large integrated circuits (VLSI). It is well known that the key points for the success of design
automation in digital systems are simple and systematic circuit design procedures and the use of standard cell circuits, labeled *gates*. In analog systems, the lack of simple design procedures and an analog gate result in the design of analog circuits still being limited to manual custom design and the initiation of analog VLSI just being started [154-160, 166]. In transconductance-C based ASP systems, transconductors can be thought of as "analog gates," because they are basic elements in ASP circuits, for instance, in filters [1, 161-166]. Furthermore, resistors, inductors, and amplifiers can be treated as "composite gates," because they all can be built from "analog gates," transconductors, and capacitors as mentioned above. On the other hand, because a transconductance-C based ASP system is actually a simulated passive RLC system, the design procedures for passive RLC systems which have been developed for over 50 years can be used with the necessary modifications in the design of transconductance-C based ASP systems [1, 23, 119, 167-173]. For example, the design procedures for LC ladder filters and cascade biquadratic filters are simple, mature, and systematic. Then, with analog gates or composite gates and the simple design procedure, computer-aided design automation for high-speed transconductance-C based APS circuits will be possible.

In the design of transconductors, linearity, defined in terms of output current related to input voltage, is a basic but important and critical parameter. The reason is that parameters of a transconductance-C based ASP circuit, for example the center or cutoff frequency of a filter, are functions of the magnitude of the transconductance, $g_m$; a constant center frequency of a filter requires a constant $g_m$ of transconductors over the whole input voltage range. Unfortunately, in practice, due to the nonlinearity of the characteristics of transistors (square-law model for MOSFET and MESFET and exponential model for bipolar), the output current of a transconductor is a linear function of input voltage only over a small range [49, 50, 64, 113, 174-177]. This precludes the use of transconductors in many potential applications, and results in considerable research having gone
into improving the linearity of transconductors over a wide input voltage range [4, 42-70, 178]. So far, the reported improvements typically are based on the differential pair circuit with current differencing for nonlinearity cancellation [4, 42-46, 56, 61-65, 67-69, 97], but the improved linearity is still limited to a relatively small input range due to the differential pair used. In Chapter II of this dissertation, two kinds of principles for the design of transconductors are proposed for improving linearity: one is based on cross-coupled pair and another is to use multiple differential pairs. Two transconductors will be built by the first principle with a CMOS cross-coupled pair and an NMOS cross-coupled pair, and they have extremely high linearity which is good for high-precision filter design [65, 64]. Also, two transconductors will be built by the second principle with current addition and current subtraction [60, 144], and they have a high linearity over a very large input range which is useful for general ASP circuit design and, particularly, for the ASP circuits operating under low power supplies [40, 179-181].

Another consideration for transconductor design is that magnitude and phase of $g_m$ must be changeable electrically to guard against poor agreement between the designed and the fabricated values of transconductors and capacitors. This disagreement usually is caused by the process tolerances or variations in operating conditions, such as temperature, power supply voltage, or aging, and results in errors in designed ASP circuits [1, 6, 182-186]. Using transconductors with variable $g_m$, these errors can be corrected by an automatic on-chip tuning scheme which is designed with the filter system and changes $g_m$ electronically and automatically. Here, the change of the magnitude of $g_m$ will be used to adjust the frequency parameters of filters, and the change of the phase will be employed to modify the quality factors [1, 187-191]. In this dissertation, all OTAs will be designed with tunable $g_m$.

Besides linearity and tuning, power supply rejection (PSR) and common-mode rejection (CMR) are two further important parameters of transconductors. Here, PSR is a
measure to show how parasitic signals (noise) couple into the signal path from power supplies and, CMR is one to show how the common-mode input signal is amplified compared to the differential input signal [41, 113-116]. Currently, the most effective method for rejecting the noise coupled from power supplies and for reducing common-mode output voltage is to use a fully-balanced circuit structure which has two exactly equal paths from inputs to outputs. When the total output voltage of this kind of circuit is defined as one output voltage referred to the other output, not to ground, the coupled noise and amplified common-mode input voltage on one output will be canceled by those on the other output [113, 192-194]. In addition, the fully-balanced circuit structure can also improve the output dynamic range and harmonic distortion [195-198]. Compared to a single-ended output circuit, no additional elements are required for a fully-balanced circuit structure; it is only necessary to break the current mirror to form two current sources. But, because the current mirror is broken, the feedback loop for the common-mode output signal is broken too. This results in an undefined common-mode output voltage and unstable DC bias voltages. The most common method used to stabilize the common-mode output voltage is to add an additional feedback circuit, called the common-mode feedback (CMF) loop, for the common-mode output signal [93, 96, 199-201]. The requirements for a CMF circuit are that, for common-mode output signals, it acts as two resistors parallel to the two outputs and ground with a small impedance and, for differential output signals, it ideally is an open circuit. In practice, to design a CMF circuit without distorting the differential signal is a challenging task [195, 201-203]. At the time of the beginning of this thesis, many CMF circuits were reported for use in op-amps [94-96, 192-198], and only few were used in CMOS operational transconductance amplifiers (OTAs). Although the circuit structure of an OTA is similar to that of an op-amp, to employ the CMF circuits used in op-amps directly for OTAs, generally, is impossible, because tuning is always required in an OTA, but not in an op-amp [195, 202, 204]. So,
there is considerable effort to modify the existing CMF circuits for OTAs and, so far, several such circuits have been successfully used. But, either the designed circuits are too complicated or the design constraints are too severe so that the requirements on fabrication processes are very critical [102, 127, 133, 135]. In Chapter III of this dissertation, a careful and systematical investigation of the problem of the inter-relation between the CMF and tuning circuits will be performed, and three different CMF circuits suitable for OTAs will be proposed.

Besides the major considerations described above, there are also some other concerns in OTA design, such as high input and output impedances, low noise, low sensitivities to process tolerance and temperature, small chip area (simple OTA), low power dissipation, and DC level equality [1, 3, 97, 144]. The last property means that the DC voltage level shift from input to output should be nearly equal to zero so that direct coupling is possible. All these considerations will be discussed in Chapter II.

In Chapter IV, two fully-balanced continuous-time transconductance-C filters will be designed with the new OTAs; one is designed in 2 μm CMOS technology as a simulated LC ladder filter with a cut-off frequency of 20 MHz [205], and the other one is designed in 3.5 μm CMOS technology as a cascade of biquadratic sections with linear phase up to 4 MHz [134]. The principle of OTA-C filter design will be discussed and the design procedures for both filters will be described step by step to show the simple and systematic design procedure. From these two design examples, the early assertion is verified: "it is not only easy to design transconductance-C filters within the megahertz range, but it also results in a simple design procedure for design automation."

It is well-known that in filter design procedures, ideally, perfect OTAs, i.e. voltage-controlled current sources, are required [138-141]. But, in practice, the parasitics cannot be totally avoided in OTA design [55, 66, 176, 246, 247]. In the lower frequency range until several hundred kilohertz, these parasitics are not very important because they
are usually much smaller than circuit capacitors which are inversely proportional to the bandwidth of a filter. When filters operate in the frequency range of several ten or hundred megahertz, parasitics become so important that even small stray components will limit frequency performance and vary the required parameters of filters [1, 6, 119, 144, 205]. For example, for given values of \( g_m \), the smaller the circuit capacitances are, the higher is the frequency response of a filter. But, the circuit capacitors should be maintained larger than the parasitic capacitances by a factor three to five so that the effect of unpredictable parasitic capacitances will be small [116, 158]. On the other hand, if large circuit capacitors are used, the frequency response of filters are reduced. So, in Chapter IV of this dissertation, the impact of the parasitic parameters on the behavior of filters will be investigated systematically to find out which parasitics are more important for the filter design and which ones less [131]. Finally, a design procedure will be developed which will help to choose the proper values for OTAs and capacitors so that the designed filters have a better performance and smaller process tolerances. Of course, this good performance will be obtained over the complete tuning range.

In order to justify another statement: "continuous-time filters are the only candidates able to use the whole frequency range that an available technology can offer," the OTA should be designed in a different technology to obtain the higher frequency response. Currently, two technologies can reach several tens of gigahertz in frequency range: GaAs MESFETs and silicon bipolar transistors. A GaAs MESFET is similar to a CMOS MOSFET in the physical device structure, and its fast speed is due to the material, Gallium and Arsenide, used [206-208]. A bipolar transistor uses the same material as CMOS MOSFET, silicon, and its fast speed comes from the physical device structure. In this dissertation, we decide to choose GaAs technologies for the OTA design for two purposes. First, when we started to design the GaAs OTA, there were no GaAs OTAs built yet but, there were many bipolar OTAs already. We wanted to design the first
GaAs OTA in this field to see whether this technology is suitable for OTA-based applications. Second, because the GaAs material is totally different from silicon, there are many problems that needed to be solved.

Although many GaAs analog circuits were designed already [208-216] at the time the GaAs OTA was designed, there were no OTAs built in this technology reported yet in the literature [217, 218]. The following problems we met in the design: (1) Because of the higher surface states of gallium arsenide material, the oxidation is not stable. So, the GaAs transistors usually is not built with oxide layer for gate isolation. Then, no MOSFET can be built in GaAs technology [219-221] which results in the input range of GaAs MESFET being strongly limited by the Schottky diodes [174, 222-224]. (2) Due to the higher mobility of gallium arsenide material and the short channel used (usually less than 1 μm), the output impedance of GaAs MESFETs is very low [225-230]. (3) In GaAs MESFET, only N-channel transistors were available when we started this thesis. Currently, to build P-channel transistors is possible, but is very complicated. So, in thesis, the design of a GaAs OTA is based on N-channel transistors only, which it is still very leads to difficulties in design [1, 144, 231].

For problem (1) listed above, nothing can be improved through circuit design; the only choice is to use differential pairs to double the dynamic input range. If the threshold voltage of the Schottky diode is e.g. 0.7 V as a typical value, the differential input range is 1.4 V. For a filter operating at very high frequencies, this input range is generally acceptable. For problem (2), there exist many techniques for enhancement of output impedances [41, 232-237], for example the cascode stage is one of the most popular techniques. Unfortunately, this technique cannot be used in GaAs OTAs because it requires complementary devices. Then, the only candidate is the "self-bootstrapping" technique, because it can be used with just N-channel devices [238-240]. The problem of "self-bootstrapping" is that, while the output impedance of GaAs OTA is increased by using
this technique, a large output capacitance is also introduced, which decreases the output impedance at high frequencies [97, 241-243]. In Chapter V, a new technique will be proposed, which increases the output impedance with only a small parasitic capacitance. The last problem mentioned above is tuning which is usually implemented by adjusting the DC tail current sources (N or P MOSFETs) with top current mirrors (P or N MOSFETs) [3, 217, 244, 245]. Obviously, this requires complementary devices which are not available in GaAs technology. In addition, Schottky diodes of MESFETs also limit the tuning range [1, 144]. So, in Chapter V, a new technique is introduced for tuning with only N-channel transistors and a digitally programmable tuning scheme is proposed for increasing the tuning range. Note that here the digital tuning signal goes only through the control paths and does not sample the analog signal. Using the above-mentioned techniques, a GaAs OTA with 7 GHz bandwidth, over 100 kΩ low-frequency output impedance, and combined digital/analog tunability is designed in 1 μm depletion-mode MESFET ($f_T = 16$ GHz) technology [97].

As mentioned early, an inductor is a key element in ASP circuits, which can be designed by OTAs and capacitance. So, in Chapter V, a GaAs gyrator-C inductor is designed. The difficulties for the design of an inductor with very-high-frequency response are the same as those for the design of high-frequency filters: the resonant frequency of an inductor is limited by the parasitics. In order to reduce these parasitics, the relationships between the inductor’s parameters and the OTA’s parameters are investigated so that it can be determined which OTA parameters are important for the design of an inductor and which ones not. Based on these decisions, a design procedure is easy to set and a GaAs inductor with near 1 GHz resonant frequency and very small parasitics is designed. With the GaAs OTAs and GaAs inductor, a high-order OTA-C filter with 200 MHz frequency response is designed.

In Chapter VI, the layout of one of the CMOS OTAs described above and of the
corresponding filter, designed by using the Magic layout graphics tool [248] in the VLSI lab in the EE department at PSU, will be presented. The IC processes will be introduced and the tools will be briefly described. The main design considerations for layout are small parasitic resistance, small parasitic capacitance, small variation of device size, small chip area, and small crosstalk. Electromigration, self heating, and hot electron problems are also considered [158, 205, 249]. A guard ring is placed around the whole N-well to prevent CMOS latch-up, and an ohmic contact is placed around each OTA for isolation [250, 251]. For design automation considerations, the layout shape of OTA is designed near square and the pins are arranged carefully so that the shortest interconnections are obtained. The circuit capacitors are built by Poly 2 over Poly 1, so that small area is required and the process tolerance is minimal [252]. The only drawback of building the capacitors in this way is that, if a floating capacitor is built, a relatively large parasitic capacitance exists between Poly 1 and ground. In order to avoid this problem, the pre-distortion method is used. This method is also used for reducing the effects of parasitic capacitances [1, 205]. The performance verification is carried out by simulating the SPICE files extracted from the layouts which, therefore, contain all information about interconnection parasitics, such as overlap and coupling capacitances [253]. The designed layouts are sent to MOSIS and fabricated by an analog process. The layouts for the other CMOS OTA and filter, fabricated by Allegro Microsystem, Inc., and one for GaAs OTA, fabricated by TriQuint Semiconductor, Inc., are designed in those companies using their resources because of confidential process technologies. The fabrication was free of charge in trade for the companies' right to use the OTAs and filters for their products; this shows that the research described here meets industrial interests.

The performance of all chips was measured in the analog testing lab in the EE Department of PSU. The main difficulty for measurements are the high-frequency parasitic effects, for example, even a short length of wire will generate a parasitic pole and
maybe cause oscillations [254]. So, in the measurements, a careful arrangement for interconnections is made and the wires used are as short as possible. The other difficulty is caused by the properties of OTAs: current output and high output impedances [205]. As mentioned earlier, matching interfaces between the outputs of OTAs and equipment is difficult. Also, the lack of test setups for the chips from MOSIS and Allegro Microsystem causes trouble. Finally, all problems were overcome and the measured results show that all circuits are functional. Except for the GaAs OTA, the results of all other OTAs and filters, are very close to the simulation results, which proves the correctness of both circuit and layout designs. For the GaAs OTA, the linearity is fine, the sinusoidal response looks good, but a good frequency response is measured only to 100 MHz and not to 7 GHz because of the lack of high-frequency equipment in the EE department. A detailed description of the measurements will be given in Chapter VII.

Conclusions are drawn in Chapter VIII.
CHAPTER II
HIGH-LINEARITY CMOS OTAs

II.1. BASIC PRINCIPLES

II.1.1. Transconductor as a Basic Building Block

As mentioned in introduction, a transconductance amplifier (transconductor) is a basic element in transconductance-C ASP circuits and, ideally, is a voltage-controlled current source as shown in Fig. 1. If nodes 1 and 2 in Fig. 1 are considered as the controlling nodes and nodes 3 and 4 as the controlled nodes, the output current, \( I_o \), is determined by the input voltage \( V_i \), \( I_o = f(V_i) \); the slope of this relation is the transconductance, \( g_m \), defined as:

\[
g_m = \frac{\partial I_o}{\partial V_i}
\]  

(II.1)

The unit of transconductance is ampere/volt which is called Siemens.

Using transconductors, a resistor with positive resistance \( (1/g_m) \) can be easily built by simply connecting nodes 1 and 3 and nodes 2 and 4 together, respectively (see Fig. 2a). If node 1 is connected to node 4 and node 2 to node 3 as shown in Fig. 2b, a negative resistor is implemented. Also, a voltage amplifier can be built by using a
transconductor and a resistor as shown in Fig. 3a with the gain being \(-g_m \cdot R\); note that here the resistor R can be replaced by a transconductor, also, so that a voltage amplifier can be built by two transconductors with the different transconductances. If these two transconductors are identical, a unit gain buffer is built. In Fig. 3b, an integrator, \(v_o = \frac{-g_m}{C} \int v_i \, dt\), is built by a transconductor terminated with a capacitor. Furthermore, a

Figure 2. Resistors built by transconductor.

Figure 3. A voltage amplifier (a) and an integrator (b) built by transconductor.
gyrator can be simulated by two transconductors as shown in Fig. 4a, and, therefore, an inductor can be implemented in gyrator-C configuration (see Fig. 4b) with the inductance $C/g_m^2$. Generally, all electrical elements used in ASP circuits can be modeled, simulated, and implemented by transconductance elements and grounded capacitors. Because a capacitor can be easily built in IC using silicon-oxidation, a direct simulation of passive ASP circuits in fully-integrated form is possible and transconductors become very important elements.

![Diagram of gyrator and inductor built by transconductors](image)

**Figure 4.** A gyrator (a) and an inductor (b) built by transconductors.

**II.1.2. Transconductor Based on a Single MOSFET**

A Metal-Oxide-Silicon Field-Effect-Transistor (MOSFET) is a four-terminal electronic element; its symbol is shown in Fig. 5a and its equivalent small-signal model in Fig. 5b. Following the general notation, $V_{ij}$ represents the voltage between terminals $i$ and $j$ and $I_i$ is used for the current through terminal $i$. Using the square-law model [49, 174, 251] with the condition $V_{sb} = 0$ (no body-effect), the drain current, $I_d$, of a MOSFET is determined by the voltages $V_{gs}$ and $V_{ds}$, and can be written as
Figure 5. A symbol (a) and a small-signal AC model (b) of a MOSFET. "D" represents the drain terminal of the MOSFET, "G" the gate, "S" the source, and "B" the bulk. \( C_{gd} \) is the gate-drain capacitor, \( C_{gs} \) the gate-source capacitor, \( C_{bd} \) the bulk-drain capacitor, and \( C_{bs} \) the bulk-source capacitor. \( R_D \) is the bulk resistor of the drain diffusion, \( R_S \) the bulk resistor of the source diffusion, and \( r_{ds} \) the small-signal output resistor between source-drain.

\[
I_d = \begin{cases} 
0 & \text{if } V_{gs} \leq V_T \\
 k[2(V_{gs} - V_T)V_{ds} - V_{ds}^2](1 + \lambda V_{ds}) & \text{if } V_T \leq V_{gs} \text{ } \& \text{ } V_{ds} \leq V_{ds, sat} \\
 k(V_{gs} - V_T)^2(1 + \lambda V_{ds}) & \text{if } V_T \leq V_{gs} \text{ } \& \text{ } V_{ds, sat} \leq V_{ds} 
\end{cases} \quad (\text{II.2})
\]

where

\[
k = \frac{\mu C_{ox} W}{2 L} \quad (\text{II.3a})
\]

is called transconductance parameter and

\[
V_{ds, sat} = V_{gs} - V_T \quad (\text{II.3b})
\]
is saturation voltage. $\mu$ is the mobility of electrons or holes, $C_{ox}$ is gate oxide capacitance per unit area, $W$ and $L$ are channel width and length, respectively, $V_T$ is the threshold voltage, and $\lambda$ is the channel-length modulation parameter.

In Eq. (II.2), the range $\{V_{gs} \leq V_T\}$ is called cut-off region, the range $\{V_T \leq V_{gs} \& V_{ds} \leq V_{ds, sat}\}$ is the triode region, and the range $\{V_T \leq V_{gs} \& V_{ds, sat} \leq V_{ds}\}$ is the saturation region. Of these regions, generally, a MOS transistor is biased in the saturation region with the following reasons: (1) A MOS transistor is easier to bias in the saturation region than the triode region due to the limitation $V_{ds} \leq V_{ds, sat}$ on the triode region. (2) In the saturation region, it can be seen from Eq. (II.2) that $I_d$ is large and is almost independent of $V_{ds}$ because $\lambda$ is, generally, small [174, 251]. (3) The output impedance, $r_{ds}$, is near constant with a large value determined by the following equation:

$$r_{ds} = \frac{\partial V_{ds}}{\partial I_d} = \frac{1}{\lambda k (V_{gs} - V_T)^2} = \frac{1 + \lambda V_{ds}}{\lambda I_d} \quad \text{(II.4)}$$

From Eq. (II.2), it also can be observed that, when a MOSFET operates in saturation with its gate and source being treated as input terminals and its drain and source as output terminals, it is a voltage-controlled current source, or a transconductor, with the transconductance being

$$g_m = \frac{\partial I_d}{\partial V_{gs}} = 2k (V_{gs} - V_T) (1 + \lambda V_{ds}) = \frac{2I_d}{V_{gs} - V_T} \quad \text{(II.5)}$$

If a resistor is implemented by a transconductor as mentioned in Section II.1.1, the resistance, $R$, is inversely proportional to $g_m$. Then, a simulated resistor with constant $R$ requires a constant $g_m$. Also, in a gyrator-C inductor, $L = C/g_m^2$ results in the requirement of a constant $g_m$. In a transconductance-C filter, the cutoff or center frequency is determined by $g_m/C$, so a constant $g_m$ is needed again. Generally, in a transconductance-C ASP system, to use a transconductor with a constant $g_m$ is a basic
requirement. Of course, it is well understood that this requirement, usually, is only satisfied within some limits acceptable to the applications; for example, for some applications $g_m$ is required to be a constant only within $|V| \leq 1 \text{ V}$.

It can be observed that, in Eq. (II.5), $g_m$ is a function of $V_{gs}$ and $V_{ds}$. Generally, $\lambda$ is relatively small, for example, $\lambda = 0.1 \text{ V}^{-1}$ for the channel length $L = 3 \text{ \mu m}$. So, the effect of $V_{ds}$ is of second-order and, usually, can be ignored in basic analysis and hand calculation. But, because $g_m$ is a linear function of $V_{GS}$ with a DC offset $-2kV_T$, the effect of $V_{GS}$ can not be ignored and should be minimized.

The non-constant $g_m$ of a single-MOSFET transconductor is caused by the non-linearity of V-I characteristics of MOSFETs; it can be seen from Eq. (II.2) that, when a MOSFET operates in saturation, the output current $I_d$ is a square function, instead of a linear function, of $V_{gs}$. If the output current of a transconductor can be designed to be a linear function of input voltage, a constant $g_m$ can be obtained. So, in the remainder of this chapter, the discussion will be focussed on how to design a transconductor with a linear characteristic between output current and input voltage.

II.1.3. Transconductor Based on a Differential-Pair

Figure 6 shows a well-known circuit structure called "unsymmetrical" differential pair [113] with a current mirror which reflects current in transistor M1 to the current in transistor M2 so that the output current $I_o$ is equal to the difference of the two currents in M1 and M2. The DC current source, $(n+1)I$, is called the tail current. When $n = 1$, it is a matched differential pair and its output current $I_o$ can be calculated as follows:

$$I_o = \begin{cases} 
I & V_i \geq \left(\frac{2I}{k}\right)^{1/2} \\
2\sqrt{kT} V_i (1 - \frac{kV_i^2}{4I})^{1/2} & |V| \leq \left(\frac{2I}{k}\right)^{1/2} \\
-I & V_i \leq -\left(\frac{2I}{k}\right)^{1/2}
\end{cases}$$

(II.6a)
The corresponding transconductance is

\[ g_m = \sqrt{8kT} \left( [2(1 - \frac{kV_i^2}{4I})] - [2(1 - \frac{kV_i^2}{4I})]^{-1/2} \right) \quad \text{for } |V_i| \leq \left( \frac{2I}{k} \right)^{1/2} \]  

and \( g_m = 0 \) for other ranges. In Eq. (II.6), \( I \) is the DC bias current, \( V_i \) the differential input voltage, and all of other parameters were defined in Eq. (II.2).

It can be observed from Eq. (II.6) that, if \( V_i \) satisfies the condition

\[ \left( \frac{4I}{k} \right)^{1/2} \ll V_i \ll \left( \frac{4I}{k} \right)^{1/2}, \]  

\( I_o \) is a linear function of \( V_i \) with a constant \( g_m \)

\[ I_o = 2\sqrt{kI} V_i \quad \text{and} \quad g_m = 2\sqrt{kI} \]  

Generally, the range defined by Eq. (II.7) is too small so that the applications of a simple differential pair transconductor are very limited. For example, if the nonlinearity error of \( I_o \) is required to be less than one percent, \( V_i \) should satisfy the condition
$V_i \leq 0.14\sqrt{4I/k}$ [45], typically $0.07$ V for $I = 200\mu A$ with $W/L = 100/2$ and $\mu n C_{ox}/2 = 66\mu A/V^2$, which is not acceptable for many applications.

By comparing the single-MOSFET and the differential pair transconductors, it can be observed that the input linear range of latter is larger than that of former, but the number of transistors used in former is only 40 percent of that used in latter. The conclusion obtained from this analysis is that, simple transconductors tend to have a smaller linear input range than more complicated transconductors.

**II.1.4. General Overview**

In recent years, there were many research efforts on how to design a linear transconductor with a large input range [4, 42-70, 118, 123]. Generally, the methods used for improving linearity can be divided into three different categories. In the first category, transconductors are built by differential pairs with source degeneration, such as those presented in References 97, 102, and 127. The principle used for improving the linearity of these kinds of transconductors is local negative feedback with feedback resistors so that the output current of the transconductor can be constrained by the feedback resistors which possess a linear relationship between voltage and current [113-116]. The advantage of this technique is that the circuitry is simple. The disadvantage is that although a strong feedback caused by large resistance can result in a large linear input range, it leads to a small transconductance. This means that a large transconductance only can be obtained in a small linear input range, say, less than 1 V for a differential input. So, the application of this technique is limited.

In the second category, transconductors are built with MOSFETs operating in triode range [54, 59, 255, 256]. It can be seen that from Eq. (II.2), when a MOSFET operates in triode range, the transconductance is

$$g_m = 2kV_{ds}(1 + \lambda V_{ds})$$  \hspace{1cm} (II.9)
In Eq. (II.9), \( g_m \) is independent of \( V_{gs} \), or of input voltage \( V_i \) if \( V_i \) is set to be equal to \( V_{gs} \). But, for a constant \( g_m \), \( V_{ds} \) should be fixed when the output current \( I_o \) changes; this is not an easy task. Furthermore, in order to guarantee that the MOSFET works in the triode range, \( V_{ds} \leq V_{ds, sat} \) must always be satisfied -- another difficult requirement. Although the techniques used in output impedance enhancement for op-amps [15, 41, 116] or for current sources [257, 258] can be used to help partially solve the above problem, the difficulties of building this kind of transconductor resulted, so far, in only few transconductors designed in this category.

Most transconductors of the third category published in the literature or used in commercial products were designed with MOSFETs operating in saturation, so that the above limitations can be ignored. The problem for this kind of transconductor is that the square-law characteristic of MOSFETs results in a nonlinear relationship between \( I_d \) and \( V_{gs} \), which further causes a non-constant \( g_m \). So far, many researchers focused on how to overcome this problem and the method used is to cancel the nonlinearity of MOSFETs by subtracting two currents with the forms \( k(a + b)^2 \) and \( k(a - b)^2 \) [42, 49, 51, 56, 62-68, 259]. With different techniques for generating the currents of the forms \( k(a + b)^2 \) and \( k(a - b)^2 \) for current subtraction, many transconductance circuits were proposed with a constant \( g_m \) in an acceptably linear range. The limitations on the circuits designed by this method, generally, are: (1) since these circuits are still built in a differential pair structure for considerations of PSRR, CMRR, and noise, the linearity of the transconductor is limited and the achieved linear input range is relatively small. (2) It is well-known that the square-law characteristic of MOSFET is only an approximate model for explaining the basic behavior of a MOSFET and to calculate the parameters of a MOSFET by hand; thus the high-order effects of a MOSFET, such as body-effect, channel-length modulation, high-injection, etc., are not included in this model. Therefore, the experimental results of a designed transconductor will not be as good as the simulation results due to
the distortion caused by these high-order effects and variations of temperature, power supplies, process tolerance, etc.

In the remaining sections of this chapter, focusing on the problems in the third category, several transconductors will be proposed. In Section II.2, a transconductor will be built with both differential input and output, but not in differential pair structure. Then, very high linearity can be achieved. In Section II.3, a transconductor will be designed in differential pair structure, but the method used for nonlinearity cancellation is not based on the method described above. A totally new idea will be proposed so that a very large linear input range can be obtained with very good linearity and the achieved linearity is not based on the square-law model.

II.2. TRANSCONDUCTOR BASED ON CROSS-COUPLED PAIRS

As mentioned in the last section, the linearity of an OTA (Operational Transconductance Amplifier) built in differential pair structure is limited and the achieved linear input range is relatively small. The reason is that the drain currents of transistors M1 and M2 in Fig. 6 are limited by the condition $I_{M1} + I_{M2} = 2I$; thus the currents $I_{M1}$ and $I_{M2}$ are not independent of each other. This results in $I_{M1}$ and $I_{M2}$ never being of the forms $(a + b)^2$ and $(a - b)^2$, respectively, at the same time. In this section, two OTAs will be built in cross-coupled pair structure instead of differential pair structure, so that the above constraints can be avoided. The design considerations on PSRR, CMRR, and noise effects will be compared to that of the differential pair.

II.2.1. CMOS Cross-Coupled Pairs

A. Basic Principle. In Fig. 7, two CMOS pairs (M3, M4) and (M5, M6) are connected in a cross-coupled configuration [65]. Usually, in a typical cross-coupled configuration, the gates of M4 and M5 are directly tied to those of M6 and M3, respectively [113], but, in Fig. 7, former is connected to latter through the transistor M1 or M2.
Because M1 and M2 work as source-followers biased by the DC currents $I_B$, the voltage drop between the gate of M3 and the gate of M5 is a constant [56]. So, the gates of M3 and M5 are tied together through a constant DC voltage, $V_{gs,M1}$. Using the square-law model in Eq. (II.2) and assuming that all transistors in Fig. 7 work in the saturation region, the currents in M3, M4, M5, and M6 can be written as follows:

\[
I_1 = I_{M3} = k_n (V_{gs,M3} - V_{Th})^2
\]

\[
= I_{M4} = k_p (V_{gs,M4} - V_{Tp})^2
\]

\[
I_2 = I_{M6} = k_n (V_{gs,M6} - V_{Th})^2
\]

\[
= I_{M5} = k_p (V_{gs,M5} - V_{Tp})^2
\]

where the subscripts $n$ and $p$ represent the N- and P-channel MOSFET, respectively, and

Figure 7. Linear transconductor circuit based on CMOS cross-coupled pairs.
the channel length modulation is ignored. Considering the following relationship among
$V_1, V_2, V_M$, and $V_N$, with the nodes $N$ and $M$ shown in Fig. 7,

\[ V_1 - V_{gx,M3} - V_{gx,M4} - V_N = 0 \]  
\[ V_2 - V_{gx,M6} - V_{gx,M5} - V_M = 0 \]

$I_1$ and $I_2$ can be obtained through simple calculations:

\[ I_1 = k_{\text{eff}} (V_1 - V_N - V_{TE})^2 \]  
\[ I_2 = k_{\text{eff}} (V_2 - V_M - V_{TE})^2 \]

where

\[ k_{\text{eff}} = \frac{k_n k_p}{(\sqrt{k_n} + \sqrt{k_p})^2} \]

and

\[ V_{TE} = V_{Tn} + V_{Tp} \]

Furthermore, considering the following relationships

\[ V_1 - V_2 = V_M - V_N = V_d \]  
\[ V_1 - V_M = V_2 - V_N = V_B \]

where $V_d = V_1 - V_2$ is the differential input voltage, and $V_B$ is defined as

\[ V_B = \sqrt{f_B/k_n} + V_{Tn} \]

$I_1$ and $I_2$ can be expressed in terms of the voltages $V_d$ and $V_B$ as

\[ I_1 = k_{\text{eff}} (V_d + (V_B - V_{TE}))^2 \]  
\[ I_2 = k_{\text{eff}} (V_d - (V_B - V_{TE}))^2 \]

It can be observed, in Eq. (II.16), that $I_1$ has the form $(a + b)^2$ and $I_2$ has the form $(a - b)^2$. Then, using a current mirror as shown in Fig. 6, the differential output
current can be obtained as:

\[ I_{\text{out}} = I_1 - I_2 \]

\[ = 4k_{\text{eff}}(V_B - V_{T_E})V_d \]

\[ = 4k_{\text{eff}}(\sqrt{I_B/k_n} - V_{T_P})V_d \]  \hspace{1cm} (II.17)

In Eq. (II.17), \( k_{\text{eff}}, V_{T_n}, \) and \( V_{T_P} \) are only process dependent parameters, therefore, they can be treated as constant for a given fabrication process. \( I_B \) is also a constant, set by the designer. Therefore, Eq. (II.17) shows that the proposed OTA circuit in Fig. 7 has perfect linearity with a constant transconductance \( g_m = 4k_{\text{eff}}[\sqrt{(I_B/k_n)} - V_{T_P}] \).

As discussed in the introduction, the magnitude of the output current of an OTA should be electrically changeable to guard against poor agreement between the designed and the fabricated values of OTAs. From Eq. (II.17), it can be seen that the output current, \( I_{\text{out}} \), can be changed by varying the DC bias current \( I_B \) so tuning can be implemented in a filter built by this OTA.

Besides linearity, linear input range is also an important parameter of OTAs. For the OTA in Fig. 7, there are two constrains in the transistors: First, all transistors should be turned on, which results in the following equation (II.18a). Second, the assumption we made for the above analysis was that all transistors should be operated in saturation; this lead to equation (II.18b) as follows:

\[ -[\sqrt{(I_B/k_n)} - V_{T_P}] = -(V_B - V_{T_E}) \leq V_d \leq (V_B - V_{T_E}) = [\sqrt{(I_B/k_n)} - V_{T_P}] \]  \hspace{1cm} (II.18a)

\[ -V_{cc} + V_B + V_{ds,sat} \leq V_d/2 = V_{1,2} \leq V_{cc} - V_B - V_{ds,sat} \quad V_B \geq 2V_{T_n} \]  \hspace{1cm} (II.18b)

\[ -V_{cc} + V_B + V_{T_P} \leq V_d/2 \leq V_{cc} - V_B - V_{T_n} \quad V_B \leq 2V_{T_n} \]  \hspace{1cm} (II.18c)

Whether the input range is determined by Eq. (II.18a) or (II.18b) depends on which range is smaller; this means that the maximum range is obtained if the two ranges
in Eq. (II.18) are equal to each other. It also can be seen from Eq. (II.18) that if a large
$V_B$ is used, the range in (II.18a) is improved but in Eq. (II.18b) is reduced, and vice
versa. So, $V_B$ is a very important parameter to determine the input range. Based on the
condition for maximum range and the relationship $V_{ds,sat} = V_B - V_{Tn}$ for the transistor
which will be used to replace the current source $I_B$, the following equation is obtained:

$$V_B - V_{Tn} = 2(V_{cc} - 2V_B + V_{Tn})$$

This results in

$$V_B = \frac{(2V_{cc} + 3V_{Tn} + V_{Tp})}{5}$$

Then, the maximum range is determined by the following equation

$$\frac{(2V_{cc} - 2V_{Tn} - 4V_{Tp})}{5} \leq V_d \leq \frac{(2V_{cc} - 2V_{Tn} - 4V_{Tp})}{5}$$

If $V_{cc} = 5$ V and $V_{Tn} = V_{Tp} = 0.7$ V are chosen, the differential input range, $V_d$, is
near 1.16 V with $V_B = 2.56$ V. Note that the equivalent input range for $V_1$ is only 0.58 V,
which is not enough for special ASP applications. Below, we will introduce two methods
for improvement.

Please note that the $V_B$ used in Eq. (II.18) is $V_{gs,M1}$ and the $V_B$ in
$V_{ds,sat} = V_B - V_{Tn}$ is $V_{gs,Ms}$, where $Ms$ is the label for the transistor used to replace the
current source $I_B$. If $V_{gs,M1} = mV_{gs,Ms}$ is considered, Eq. (II.18) becomes:

$$-(mV_{gs,Ms} - V_{Tn}) \leq V_d \leq (mV_{gs,Ms} - V_{Tn})$$

$$-V_{cc} + (m + 1)V_{gs,Ms} - V_{Tn} \leq V_d/2 \leq V_{cc} - (m + 1)V_{gs,Ms} + V_{Tn}$$

As done before, setting two ranges in (II.20) equal results in

$$V_{gs,Ms} = \frac{2(V_{cc} + 2V_{Tn,p})}{2 + 3m}$$

$$V_d = \frac{2m(V_{cc} + 2V_{Tn,p})}{2 + 3m} - 2V_{Tn,p}$$
Since

\[
\frac{\partial V_d}{\partial m} = \frac{4(V_{cc} + 2V_{T_{p,n}})}{(2 + 3m)^2} \geq 0,
\]

(II.22)

\( V_d \) is a monotonically increasing function of \( m \). This means the input range enlarges with \( m \) increasing. For Eq. (II.19), \( m = 1 \) was chosen; the input range, therefore, was small.

From Eq. (II.21a), it is easy to obtain that

\[
m = \frac{2(V_{cc} + 2V_{T_{p,n}})}{3V_{gs,Ms}} - \frac{2}{3}
\]

(II.23)

Equation (II.23) shows that \( m \) is an inverse function of \( V_{gs,Ms} \). Therefore, a minimum \( V_{gs,Ms} \) leads to a maximum \( m \) which results in a maximum \( V_d \). Due to the constraint from Eq. (II.18c), the minimum \( V_{gs,Ms} \) is equal to \( 2V_{T_{p,n}} \), or 1.4 V, then \( m = 2.38 \) for \( V_{cc} = 5 \) V and \( V_{T_{p,n}} = 0.7 \) V. This leads to \( V_d = 1.93 \) V which is 66% larger than the input range we got before, 1.16 V.

\( m = 2.38 \) can be easily obtained by choosing the ratio of the widths of Ms to M1, or vice versa. For example, if the width of M1, \( W_{M1} \), is selected, the width of Ms, \( W_{Ms} \), can be obtained from the following equation:

\[
W_{Ms} = W_{M1} \frac{(mV_{gs,Ms} - V_{Tn})^2}{(V_{gs,Ms} - V_{Tn})^2}
\]

(II.24)

So, the advantage of this improvement is that no additional circuit is needed. But, the disadvantage is that input range is still relatively small.

The second improvement is to change the OTA circuit in Fig. 7 slightly. Fig. 8 shows a new OTA [65] which, compared to the OTA in Fig. 7, contains only two additional transistors, M3 and M4 (please note that the numbering of the transistors in Figs. 7 and 8 is different) and two current sources \( I_{C2} \). With M3 and M4 and the current sources, the input range of the OTA in Fig. 8, becomes
Figure. 8. Modified arrangement of the transconductor circuit of Fig. 7.

\[-\sqrt{I_C/k_{eff}} = -(V_C - V_{TE}) \leq V_d \leq (V_C - V_{TE}) = \sqrt{I_C/k_{eff}}\]  

(II.25a)

\[-V_{CC} + V_C + V_{ds,\text{sat}} \leq V_d/2 = V_{1,2} \leq V_{CC} - V_{C1} - V_{ds,\text{sat}}\]  

(II.25b)

where $I_C = I_{C1} = I_{C2}$ with $I_{C1}$ and $I_{C2}$ shown in Fig. 8. $V_C$ is defined as the voltage drop between the nodes $V_P$ and $V_M$ (see Fig. 8) and determined by the following equation:

\[V_C = \sqrt{I_{C1}/k_n} + \sqrt{I_{C2}/k_p} + V_{TE} = \sqrt{I_C/k_{eff}} + V_{TE}\]  

(II.26a)

Further

\[V_{C1} = \sqrt{I_C/k_n} + V_{Tn}\]  

(II.26b)

Obviously, as discussed before, the maximum range is obtained only if the two ranges in Eq. (II.25) are equal. From this condition, the input range for the OTA in Fig. 8 can be calculated as

\[-2(V_{CC} - V_{Tn,p})/3 \leq V_d \leq 2(V_{CC} - V_{Tn,p})/3\]  

(II.27)
with $V_C = 2V_{C1} = 2(V_{cc} + 2V_{Tp,n})/3$

If $V_{cc} = 5\, \text{V}$ and $V_{Tn} = V_{Tp} = 0.7\, \text{V}$ are still chosen, $V_{C1} = 2.135\, \text{V}$ and $V_d = 2.87\, \text{V}$ which is more than twice of that of OTA in Fig. 7. Note that, because $V_{C1}$ is larger than $2V_T$, the above value is valid.

It should be pointed out that the OTA in Fig. 8 has the same linearity as that in Fig. 7. A brief analysis is given below.

Since

$$V_1 - V_2 = V_M - V_N = V_P - V_Q = V_d$$

$$V_P - V_M = V_Q - V_N = V_C$$

Then,

$$I_1^* = k_{eff} (V_d + V_C - V_{TΣ})^2$$

$$I_2^* = k_{eff} (V_C - V_d - V_{TΣ})^2$$

Therefore,

$$I_{out}^* = I_1 - I_2 = 4k_{eff}(V_C - V_{TΣ}) = 4\sqrt{k_{eff} I_C} V_d$$

The advantage of this improvement is that the input range has more than doubled. The disadvantage is that an additional four transistors are used so that the silicon area is larger and power is almost doubled.

Obviously, the above two improvements can be combined together to obtain an ever larger input range. The proof is not provided here.

**B. Simulation Results.** In order to verify the above analysis, the OTA in Fig. 8 was built with current mirrors and output stages (see Fig. 9). It was simulated by TSPICE, a SPICE with Tektronix interface environment, with a semi-empirical curve-fitting (Level 3) model. The parameters of the MOSFETs used are $V_{Tn} = 0.798\, \text{V}$, $V_{Tp} = -$.
Figure 9. Complete circuit diagram of the circuit in Fig. 8.
0.914 V, $\mu_n C_{ox}/2 = 66.0 \mu A / V^2$, and $\mu_p C_{ox}/2 = 24.7 \mu A / V^2$, and the gate lengths are 36 \mu m for all transistors. The gate widths for N-channel transistors are: $W_1 = W_2 = W_5 = W_8 = 8 \mu m$, and for P-channel are: $W_3 = W_4 = W_6 = W_7 = 24 \mu m$. By choosing $V_a = -3.65$ V, $m = 2.64$ can be obtained from Eq. (II.23) and $W_{Ms} = 202 \mu m$ for N-channel transistor and $W_{Ms} = 841 \mu m$ for P-channel can be calculated from Eq. (II.24). Because the channel width for P devices is too large in practice, $W_9 = W_{10} = 200 \mu m$ and $W_{11} = W_{12} = 500 \mu m$ was used. For this choice and ±5 V power supply, the simulation results in Fig. 10 indicate that this OTA has a good linearity as expected and the differential input range is near ±3 V with less than 1% nonlinearity error.

![Figure 10. SPICE simulation results for the circuit in Fig. 9. <1> DC transfer curve of the output current <2> Linearity error expressed as a percentage of the full-scale DC current.](image)

II.2.2. NMOS Cross-Coupled Pairs

The OTA discussed in the last section has high linearity over a relatively large differential input range which may, however, not be large enough for some special applications. The main limitation there is expressed by Eq. (II.18a), due to complementary
transistors used in cross-coupled CMOS pairs. In the other words, although the different mobilities of P-channel and N-channel devices do not cause any problem in a cross-coupled CMOS pair configuration [47], the different channel widths will cause a different degree of process tolerance. For example, the P-channel devices with $W = 24 \, \mu \text{m}$ (used in the OTA in Fig. 9) will have different process tolerances with the N-channel devices with $W = 8 \, \mu \text{m}$. So, in order to avoid the above two problems, an OTA with an NMOS cross-coupled pair will be introduced in this section.

A. Basic Principle. The OTA proposed here, shown in Fig. 11 [64], contains only NMOSFETs instead of the cross-coupled CMOS pairs used in the OTA in Fig. 7. Because of this change, the principle of operation of this OTA is different in a fundamental way from that of the one in Fig. 7. For example, the current $I_1$ of the OTA in Fig. 11 equals the sum of the drain currents of M3 and M4, i.e., $I_1 = I_{d3} + I_{d4}$, but in the previous OTA, $I_1 = I_{d3} = I_{d4}$. Also, the OTA in Fig. 11 is a class A amplifier, but that in Fig. 7 is class AB.

![Figure 11. Linear transconductor circuit base on NMOS cross-coupled pairs.](image-url)
When all transistors in Fig. 11 work in saturation, the drain currents of M3, M4, M5, and M6 can be expressed as

\[
I_{d3} = k_n(V_1 - V_{ss} - V_{Trn})^2 \quad I_{d4} = k_n(V_2 - V_c - V_{Trn})^2
\]
\[
I_{d5} = k_n(V_1 - V_c - V_{Trn})^2 \quad I_{d6} = k_n(V_2 - V_{ss} - V_{Trn})^2
\]

(II.31)

where \(k_n\) and \(V_{Trn}\) were defined in Eq. (II.10). Note that, because \(M_1, M_2, M_7\) and \(M_8\) carry the same currents (source followers configuration), the relationship that \(V_1 - V_M = V_2 - V_N = V_c - V_{ss}\) was used for deriving Eq. (II.31). Note that \(V_c\) used here is defined in Fig. 11.

As done in the last section, the differential output current \(I_{out}\) can be obtained by using a current mirror:

\[
I_{out} = I_1 - I_2 = I_{d3} + I_{d4} - I_{d5} - I_{d6} = 2k_n(V_c - V_{ss})V_d
\]

(II.32)

where \(V_d\) was defined in Eq. (II.14a).

It can be seen from Eq. (II.32) that this OTA has perfect linearity with a constant transconductance,

\[
g_m = 2k_n(V_c - V_{ss}).
\]

(II.33)

Similar to the OTA introduced before, the value of \(g_m\) can be tuned by changing the DC bias voltage \(V_c\).

It can be seen that there are two advantages of this OTA over the OTA in Fig. 7. First, the same \(k_n\) used means all transistors in Fig. 11 are identical because they are all NMOSFETs. This results in easy control of process tolerances, but this cannot be true for the CMOS pairs. Second, it can be seen that, from Eq. (II.33), \(g_m\) of this OTA is independent of \(V_T\) which is also a process-dependent parameter, but \(g_m\) in either Eq. (II.17) or (II.30) is dependent on \(V_T\). Therefore, \(g_m\) of this OTA is easier to control in the face of process tolerances than a CMOS pair.
Considering linear input range, it can be seen from Fig. 11 that the only restriction is that all transistors of the OTA should be operated in the saturation region. This means that the differential input voltage must fulfill the condition

\[ |V_d| \leq 2|V_c - V_T| \]  

(II.34)

For \( V_{cc} = 5 \, V \) and \( V_T = 0.7 \, V \), in order to guarantee transistors M7 to M8 to be on, the maximum \( |V_c| \) is 4 V. This results in the maximum dynamic differential linear input range to be as large as ±6.6 V, which is larger than that of any OTAs introduced before.

In addition to the large input range, this OTA also has extremely high linearity. From Eq. (II.31), the output current with respect to the input voltages \( V_1 \), \( I_{o_1}V_1 \), or to \( V_2 \), \( I_{o_2}V_2 \), can be obtained:

\[
I_{o_1}V_1 = I_{d3} - I_{d5} = k_n(V_c - V_{ss})(2V_1 - 2V_T - V_c - V_{ss}) \quad \text{(II.35a)}
\]

\[
I_{o_2}V_2 = I_{d6} - I_{d4} = k_n(V_c - V_{ss})(2V_2 - 2V_T - V_c - V_{ss}) \quad \text{(II.35b)}
\]

From Eq. (II.35a), it can be seen that \( I_{o_1}V_1 \) is a linear function of \( V_1 \) with a DC offset; this means the first-order nonlinearity is canceled already by subtracting \( I_{d5} \) from \( I_{d3} \). A similar argument applies to \( I_{o_2}V_2 \). In addition, because \( I_{o_1}V_1 \) and \( I_{o_2}V_2 \) are exactly symmetrical in \( V_1 \) and \( V_2 \), respectively, another subtraction between \( I_{o_1}V_1 \) and \( I_{o_2}V_2 \) will not only cancel the DC offsets but also higher-order nonlinearities in \( I_{out} \) (see Eq. (II.32)). This results in an extremely high linearity which will be proven later with SPICE simulations based on Level 3 model.

B. Simulation Results. The OTA in Fig. 11 with a proper current mirror was simulated by TSPICE with Level 3 model for high-order nonlinearity considerations. The transistor model is the same as that used in the last section, and the sizes of all transistors are \( W = L = 10 \, \mu m \). \( V_{cc} = -V_{ss} = 5 \, V \) and \( V_c = -4 \, V \). The simulation results in Fig. 12a show that the linearity error of the proposed OTA is only ±0.1% over a differential input range as large as \( V_d = \pm 3.8 \, V \), and increases to only ±0.7% in \( V_d = \pm 6.4 \, V \) which is close
to the value predicted by Eq. (II.34) because there are no additional non-linear errors caused by high-order nonlinearity. Outside this range, as assumed in Eq. (II.34), the linearity becomes worse. When $6.4 \leq |V_d| \leq 7.8$ V, only M4 and M5 are off, but all other transistors are still in saturation. So, $I_{out}$ does not saturate but equals $I_{d3} - I_{d6}$. After $|V_d| \geq 7.8$ V, i.e., $|V| \geq 3.9$ V, $i = 1, 2$, $V_{gs1}$ and $V_{gs2}$ become non-constant because M7 and M8 enter the triode region; this leads to the linearity of the OTA becoming very poor. When $|V_d| \geq 8.4$ V, M3 and M6 are turned off and $I_{out}$ goes to zero gradually as can be seen from Fig. 12a.

The simulation results in Fig. 12b also show that the transconductance of this OTA can be tuned by a factor 3, with $V_c$ changing from -4 V to -2 V. Note that the linear input range decreases with decreasing $|V_c|$ consistent with Eq. (II.34).

Besides tunability, another important OTA parameter is frequency response. It can be seen from Fig. 11 that there are only four internal nodes A, B, C, and D in the OTA and even the nodes A and B are not critical because they are the outputs of source followers. Thus, the new OTA is expected to have good frequency response. In fact, SPICE simulation results (see Fig. 13) indicate that $f_{(-3dB)} = 63$ MHz for channel length $L = 10$ µm and $f_{(-3dB)} > 240$ MHz for $L = 5$ µm. If $L$ is chosen smaller, the frequency response increases further, but linearity is worse. Therefore, there exists a tradeoff between linearity and frequency response: larger channel length results in a smaller channel length modulation parameter which, in turn, leads to smaller variation of $(1 + \lambda V_{ds})$; thus, higher linearity and larger linear input range is obtained, but poorer frequency response.

II.2.3. Conclusions

In the last two sections, two OTAs were introduced and linearity and linear input range of each OTA had been analyzed in detail. Generally, they both have nice frequency responses, because each circuit contains few internal nodes; this had been discussed in
Figure 12. The simulation results for the OTA in Fig. 11. (a) DC transfer curves: <1> the total output current and <2> linearity error expressed as a percentage of the full-scale DC current $I_1 + I_2$ as in Reference [127]. (b) Tunability: <1> $V_C = -4 \text{ V}$, <2> $V_C = -3.5 \text{ V}$, <3> $V_C = -3 \text{ V}$, <4> $V_C = -2.5 \text{ V}$ and <5> $V_C = -2 \text{ V}$. 
Figure. 13. The simulation results for the frequency responses of the OTA in Fig. 11 with $L = 10 \, \mu m$ (r1) or $L = 5 \, \mu m$ (r0).

Section II.2.2. Tuning is easy to implement by varying DC bias voltages, which was also shown in Section II.2.2. Next, consider the common mode signals.

It is seen from either Fig. 7 or Fig. 11 that the circuit structure of these OTAs is not based on a differential pair, common-mode input signals, therefore, will be amplified. If two input signals are purely differential voltages with common-mode voltage $V_{cm}$, such as $V_1 = V_{cm} + V_d/2$ and $V_2 = V_{cm} - V_d/2$, the amplified $V_{cm}$ will be cancelled in Eq. (II.17) or Eq. (II.32) and the output currents of the OTAs will be still purely differential. This means the linearity of both OTAs will not be destroyed and the only problems are more power consumption and the reduced input range; the latter also happens to the OTA built by differential pairs. Note that process tolerance is not considered here, because it happens to all types of circuits, not particularly to the OTAs discussed here.

If the two input signals are not as described above, but are the totally different, the linearity of these OTAs may be poor due to the amplified common-mode signals.
existing in the output currents. In the following, the two OTAs will be analyzed separately, because they have different results.

First, consider the single input case in which one input signal is applied and another is connected to ground, such as $V_1 = V_d$ and $V_2 = 0$. Ideally, because all transistors in both OTAs are in the saturation region, the equations (II.12) and (II.35) are still true. This leads to equations (II.17) and (II.32) being true and the output currents supposedly linear with the input voltage $V_d$.

But, in practice, an analysis may be needed to determine the linearity based on the operating points because AC output current is determined by the small-signal model in which the AC signal is only valid in a small range near the DC operational point. For different DC operating points, the AC output current will be different even for the same input voltage. Note that this point cannot be seen from equations (II.10) and (II.31), because these equations are only for large signals and only for the first-order approximation model of MOSFETs. In Fig. 14a, SPICE simulation results based on a third-order model show that the currents in transistors M3 and M5 of the OTA in Fig. 11 are different and the nonlinearities cannot cancel each other.

It should be pointed out that the nonlinearity caused by DC operating points will cancel if differential inputs are applied to the OTAs. It stands to reason that the currents of all transistors M3 to M6 change with input voltage so that the nonlinearity of M3 is canceled by that of M6 (Note that they are at the same DC bias.) and the nonlinearity of M4 is canceled by that of M5.

For the OTA in Fig. 11, because $V_2 = 0$, the AC currents in the transistors M6 and M4 are zero. This results in the output current $I_{out}$ being determined by the AC currents in M3 and M5. Due to M3 and M5 operating at different operating points, such as $V_{gs3,0} = -V_{ss}$ and $V_{gs5,0} = -V_c$, the output currents $I_{d3}$ and $I_{d5}$ are different. The DC components in M3 and M5 can be cancelled by those in M4 and M6, but the AC mode
Figure 14. SPICE simulation results for the OTA in Fig. 11 with an asymmetrical input signals: $V_1 = V_{in}$ and $V_2 = 0$: (a) The drain currents of the transistors M3 and M5, id3 and id5, and (b) the total output current, id, and transconductance, dif(id).
cannot be cancelled because there are no AC components in M4 and M6. This leads to
the total output current, $I_{out}$, becoming nonlinear which has been verified by the SPICE
simulation results shown in Fig. 14b.

For the OTA in Fig. 7, the AC currents in the transistors M4 and M6 are not zero
even if $V_2 = 0$ because they are restricted by the currents in M3 and M5, thus

$$I_{M4} = I_{M3} \text{ and } I_{M6} = I_{M5} \quad (II.36a)$$

Actually, it can be seen from Fig. 7 that although both $V_2$ and $V_N$ are AC ground when
$V_2 = 0$, the voltages $V_{gs,M4}$ and $V_{gs,M6}$ are not constant because they are restricted by

$$V_{gs,M3} + V_{gs,M4} = V_d + V_B \text{ and } V_{gs,M5} + V_{gs,M6} = V_d + V_B \quad (II.36b)$$

The relationship in Eq. (II.36) has the same format as the relationship of the differential
pair; therefore, the output current of the single input for the OTA in Fig. 7 is the same as
that of the differential inputs. This means the linearity of single input will not be dis-
torted for the OTA in Fig. 7 at all. But, because M3 is NMOS and M4 is PMOS, not both
devices are NMOS in the differential pair, near two third of $V_d$ drops on $V_{gs,M4}$ and only
one third drops on $V_{gs,M3}$. This results in the linear input range being reduced. Also, as
there is no DC current source used to limit the total DC current, the power consumption
may increase. The above analysis is also proven by SPICE simulation and the results in
Fig. 15 show that the linearity of this OTA with the single input is as good as that with
the differential inputs and the input linear range is reduced and shifted to left.

Second, consider the case that two input signals having different magnitudes, for
example $V_1 = 1.5V_d$ and $V_2 = 0.5V_d$. For the OTA in Fig. 11, although the AC currents
in all transistors are not zero and the DC operating points of M3 and M6 (also M4 and
M5) are the same, the output linearity is still destroyed. It can be seen from Fig. 16a that
the AC current in M3, $I_3$, is different from that in M6, $I_6$, so the nonlinearities in $I_3$ and
$I_6$ cannot cancel each other. Although the applied voltages to M3 and M5 are the same,
the nonlinearities of $I_{M3}$ and $I_{M5}$ still cannot cancel each other because the DC operating points of M3 and M5 are different. The same situation exists between M4 and M6. Therefore, the output linearity is distorted and this conclusion is shown in Fig. 16b.

The above situation does not exist in the OTA shown in Fig. 7. The reason can be understood from the analysis to single-ended case; Eq. (II.36) is actually true for any kind of input voltages. Therefore, the applied AC voltages to M3 and M6 (also M5 and M4) are the same and the DC operating points of M3 and M6 are the same, the output linearity, therefore, is maintained as for the differential input case. Same as in the single-ended case, the linear input range will be reduced. This conclusion is shown in Fig. 17.

The final conclusion is that the linearity of the OTA shown in Fig. 7 is not sensitive to applying any kind of input signals, just as a differential pair. Although the OTA in Fig. 11 is also independent of the input wave forms in a first-order approximation but depend on it for the second-order approximation, the linearity is distorted. On the other
Figure 16. SPICE simulation results for the OTA in Fig. 11 with asymmetrical input signals: $V_1 = 1.5V_{in}$ and $V_2 = 0.5V_{in}$: (a) The drain currents of the transistors M3 to M5, $i_3$, $i_4$, $i_5$, and $i_6$, and (b) the total output current, $i_d$, and transconductance, $\text{dif}(i_d)$. 
Figure 17. The simulation results for the total output current, id, and transconductance, dif(id), of the OTA in Fig. 7 with \( V_1 = 1.5V_{in} \) and \( V_2 = 0.5V_{in} \).

hand, the linear input range for both OTAs is reduced more than for a differential pair when input signals are non-symmetrical, and the power consumption is also larger than in a differential pair because the common-mode input voltage is amplified in cross-coupled, but not in differential-pair OTAs.

To avoid the above problems, one choice is to use an input converter circuit which can convert inputs with any kind of wave forms to inputs with purely differential signals. A second choice is to build an OTA by differential pairs. The first choice will not be discussed in this dissertation and will be a further research topic. The second choice will be discussed in the next section.

II.3. TRANSCONDUCTOR BASED ON MULTIPLE DIFFERENTIAL PAIRS

As discussed before, an OTA based on differential pairs cannot have its two outputs of the forms \((a + b)^2\) and \((a - b)^2\), because the two output currents are constrained
by each other. This results in that the techniques for linearization based on 
\[(a + b)^2 - (a - b)^2\] cannot be applied to differential pairs without degrading linearity.

But, the advantages of the differential pair are so attractive that a large amount of research has been done. Another technique for linearization applied to a differential pair is negative feedback. But, methods for this technique either have poorer linearity than the technique employing \[(a + b)^2 - (a - b)^2\] or have a smaller linear input range. This is pronounced in FETs because \(g_m/I\) is much lower than it is for bipolar transistors.

In this section, a new approach, labeled \textit{current addition/subtraction}, [60, 144] is introduced for the design of a CMOS OTA with improved linearity over a very large input range. The method can be explained by the technique \[(a + b)^2 - (a - b)^2\], but the limitation of DC tail current (see Fig. 6) is reduced by using an additional differential pair. This method is based on the principle of, and is easy to explain by, negative feedback. The approach has three features: 1) the result is independent of the first-order model of the transistors, 2) the input range can be designed to be very wide, and 3) the approach is a general design method which can be applied not only to CMOS, but to other technologies as well.

II.3.1. Basic Principle

The circuit of a differential pair is shown in Fig. 6, and its output current and its transconductance for \(n = 1\) is given in Eq. (II.6). It was also mentioned there that a linear input range is defined in Eq. (II.7).

A more useful expression, compared to Eq. (II.7) can be obtained by defining a DC voltage, \(V_I\), such that

\[-V_I \leq V_i \leq V_I\] (II.37)

is the range in which the current \(I_o\) is approximately linear. For example, if the nonlinearity error of \(I_o\) is required to be less than one percent, \(V_I = 0.14\sqrt{4I/k}\) [45].
When \( n \neq 1 \), the circuit in Fig. 6 becomes a mismatched differential pair. The DC currents in the transistors M1 and M2 are \( nI \) and \( I \), respectively. Thus, current \( I_o \) is restricted to have a maximum value \( I \) and a minimum value \(-nI\), and a complete expression for \( I_o \) can be shown to be

\[
I_o = \begin{cases} 
  I & V_i \geq V_2 \\
  -2\alpha_1 kV_i^2 + \alpha_2 \sqrt{kTV_i(1 - \alpha_3 \frac{kV_i^2}{I})} & -V_1 \leq V_i \leq V_2 \\
  -nI & V_i \leq -V_1 
\end{cases} \tag{II.38a}
\]

where

\[
\alpha_1 = \frac{n(n-1)}{(n+1)^2}, \quad \alpha_2 = \frac{4n}{(n+1)}, \quad \alpha_3 = \frac{n}{(n+1)^2} \tag{II.38b}
\]

and

\[
V_1 = \left[ \frac{I}{k(n+1)} \right]^{1/2}, \quad V_2 = \left[ \frac{I}{k} \left( \frac{n+1}{n} \right) \right]^{1/2} \tag{II.38c}
\]

It can be seen from Eq. (II.38) that \( I_o \) becomes a nonsymmetrical function of \( V_i \) when \( n \neq 1 \). For \( n > 1 \), the small-signal range of \( I_o \) is shifted to the left, i.e., towards negative \( V_i \), because \( V_2 \leq V_1 \). For \( n < 1 \), the range is shifted to the right since \( V_2 \geq V_1 \). Thus, the center of the small-signal range can be shifted to the right or left of the point of origin, \( V_i = 0 \), by a distance approximately equal to

\[
V_s = \frac{1}{2} V_1 - V_2 = \frac{1}{2} \left[ \frac{I}{k(n+1)} \right]^{1/2} \left| 1 - \left( \frac{1}{n} \right)^{1/2} \right| \tag{II.39}
\]

This results in the linear range of the mismatched differential pair with \( n > 1 \) being limited now by

\[-V_s - V_i \leq V_i \leq -V_s + V_i \tag{II.40a}\]

and for \( n < 1 \) by
\[ V_s - V_l \leq V_i \leq V_s + V_l \]  
\[ (II.40b) \]

From Eq. (II.39), it can be seen that \( V_s \) is a function of \( n \), so that the position of the linear range of \( I_o \) can be controlled by the designer via choosing \( n \).

Note that, in practical designs the case \( n < 1 \) is not obtained by reducing the width of transistor M1 by a factor \( n \), but by interchanging transistors M1 and M2 in Fig. 6 with \( n > 1 \). Then, the current \( I_o \) for this case becomes

\[
I_o = \begin{cases} 
  nI & V_i \geq V_1 \\
  2\alpha_1 kV_i^2 + \alpha_2 \sqrt{k} I V_i (1 - \alpha_3 \frac{kV_i^2}{I}) & -V_2 \leq V_i \leq V_1 \\
  -I & V_i \leq -V_2 
\end{cases} \quad (II.41)
\]

with \( n > 1 \), and the center of the small-signal range, \( V_s \), is still determined by Eq. (II.39) with \( n > 1 \). But, because the width of M1 is less than that of M2, the linear range is restricted by Eq. (II.40b) instead of Eq. (II.40a); this is why the boundary values of \( I_o \) in Eq. (II.41) are interchanged from those of Eq. (II.38a). In the following, the case "\( n < 1 \)" is always understood to mean this situation with \( I_o \) being represented by Eq. (II.41) and \( n \) being larger than unity.

In order to verify the above analysis, the circuit in Fig. 6 has been simulated by SPICE for \( n > 1 \) (\( W_1 = 9W \) and \( W_2 = W \)), \( n < 1 \) (\( W_1 = W \) and \( W_2 = 9W \)), and \( n = 1 \) (\( W_1 = W_2 = 4W \)). The reason for choosing a large width for \( n = 1 \) is to keep the curve for \( n = 1 \) distinct from the other ones. The simulation results in Fig. 18 show that the curves for \( I_o \) have the shapes discussed above and that \( I_o \) and \( V_s \) are not only functions of the bias current \( I \), but also of the ratio \( n \), as expected. Note that, as discussed above, for the case "\( n = \frac{1}{9} \)" \( V_1 \) and \( V_2 \) are calculated by Eq. (II.38c) with \( n = 9 \).

Now consider adding the output currents of a differential pair with \( n = 1 \) and another one with \( n > 1 \) and \( V_s = 2V_i \); then, from Eq. (II.37) and Eq. (II.40a), the linear
range for this combination is ideally enlarged to $-3V_i \leq V_i \leq V_i$. If one more differential pair with $n < 1$ and $V_s = 2V_i$ is added to the above combination, the linear range becomes $-3V_i \leq V_i \leq 3V_i$. Extending this concept, if the output currents of one matched differential pair and $(m-1)$ distinct differential pairs with asymmetrical, equal, and nonoverlapping linear ranges are added, the linearity can be improved and the total linear range could be increased to $-mV_i \leq V_i \leq mV_i$, subject to limits imposed only by power supply and bias currents. Of course, this is only an ideal result because the nonlinear range of one differential pair tends to destroy the linear ranges of other pairs. In practice, according to the relationships among the output currents of each differential pair and the total output current, the centers and widths of each linear range and the magnitudes of each output current must be properly selected to minimize the nonlinear effects. The design considerations and trade-offs will be discussed in the following in connection with a practi-
II.3.2. Practical Design

The circuit in Fig. 19 has been designed according to the above considerations with \( m = 3 \), i.e., three differential pairs. The first mismatched pair contains the transistors M1 and M2 with its linear range shifted to the left of the origin, \( V_i = 0 \), because \( n > 1 \). The second pair (M3, M4) has its linear range shifted to the right. As mentioned above, when the currents of the two pairs are added, the nonlinear range of one pair may destroy the linear range of the other pair. To avoid this effect, the distance between the linear ranges of the two pairs is chosen so far that the nonlinear range of one pair does not overlap the linear range of the other. This can be done by choosing a large \( n \) according to Eq. (II.39). The possible problem caused by this arrangement is that there may exist a very bad linearity in the region near the origin \((V_i = 0)\) since that region contains only the nonlinear ranges of both pairs. To investigate this question, consider the sum of output currents of these two pairs, \( I_{a,1+2} \):

\[
I_{a,1+2} = \begin{cases} 
(n+1)I 
& V_i \geq V_1 \\
I + 2\alpha_1 kV_i^2 + \alpha_2 \sqrt{k} IV_i(1 - \alpha_3 \frac{kV_i^2}{I}) & V_2 \leq V_i \leq V_1 \\
2\alpha_2 \sqrt{k} IV_i(1 - \alpha_3 \frac{kV_i^2}{I}) & |V_i| \leq V_2 \\
-V_1 \leq V_i \leq -V_2 \\
-(n+1)I & V_i \leq -V_1 
\end{cases} \tag{II.42a}
\]

From Eq. (II.42a), it can be seen that, although the center range, \(|V_i| \leq V_2\), contains only the nonlinear ranges of two pairs, the linearity there is actually improved, not worse, because the nonlinearities of the two pairs cancel each other. Indeed, if \( n \gg 1, \alpha_3 \approx 0 \), then the current \( I_{a,1+2} \) in \(|V_i| \leq V_2\) becomes a linear function of \( V_i \) with transconductance \( 2\alpha_2 \sqrt{k} I \). From Eq. (II.42a), the sum of the transconductances of these two pairs can be
Figure. 19. Linear transconductance circuit based on multiple differential pairs.
calculated as

\[
g_{m1} + g_{m2} = \begin{cases} 
\frac{\alpha_2 \sqrt{8T}}{[2(1 - \alpha_1 - \alpha_2) - 2(1 - \alpha_1 - \alpha_2)]} & |V_i| \leq V_2 \\
4\alpha_1 kV_i + \frac{\alpha_2 \sqrt{8T}}{[2(1 - \alpha_1 - \alpha_2)]} & V_2 < |V_i| < V_1 \\
0 & \text{otherwise}
\end{cases}
\]

Equation (II.42b) indicates two problems: (i) Although \(g_{m1} + g_{m2}\) is almost independent of \(V_i\) in the center range as mentioned above, the maximum value of \(g_{m1} + g_{m2}\) in that range is less than that in the other range. (ii) \(g_{m1} + g_{m2}\) is approximately a linear function of \(V_i\) in the range \(V_2 < |V_i| < V_1\). In order to solve these problems, a matched differential pair should be added because its transconductance, Eq. (II.6b), is almost constant in the center range \(|V_i| \leq V_1\) and is a linearly decreasing function of \(V_i\) for \(|V_i| > V_1\).

After the current of the matched differential pair (M5 and M6 in Fig. 19) is added to \(I_{o,1+2}\), a good compensation of nonlinearity of \(g_{m1} + g_{m2}\) can be achieved for a fixed \(n\) if the height and width of the transconductance of M5 and M6, labeled \(g_{m3}\), are properly chosen.

Figure 20 shows the SPICE simulation results for the individual transconductances \(g_{m1}, g_{m2}, g_{m3}\), and \(g_{m1} + g_{m2}\) with Level 3 model parameters as used before. Here, a semi-empirical curve-fitting model is used for being able to consider high-order nonlinearities. It can be seen from Fig. 20 that, in order to fill the "hole" of \(g_{m1} + g_{m2}\) in the center range, the height of \(g_{m3}\) should be chosen by carefully selecting the widths of M5, M6, and M13 (the current source) in Fig. 19 as

\[
g_{m3}(0) = g_{m2}(V_{EM}) - [g_{m1}(0) + g_{m2}(0)] \tag{II.43a}
\]

\(V_{EM}\) is the voltage at which \(g_{m2}\) has its maximum value.

Also, in order to improve \(g_{m1} + g_{m2}\) in the other ranges, the total width of \(g_{m3}\)
should be chosen to equal $2V_{EM}$, that is

$$g_{m3}(V_i) = 0 \quad \text{for} \quad |V_i| \geq V_{EM}$$

The required "shape" of $g_{m3}$ can, in principle, be found from Eq. (II.43). But $V_{EM}$ is a solution of a sixth-order nonlinear equation and, as a result, Eq. (II.43) is difficult to solve. Also, as an important factor, even if the required channel widths of M5, M6, and M13 are calculated from Eq. (II.43), they will differ, in practice, from the actual values needed because the calculated widths are based on the first-order model of the transistors. Thus, Eq. (II.43) is only used to illustrate the existence of an optimum result but a solution will not be attempted. Rather, the method used to find an optimum solution for a fixed $n$ is to adjust the shapes of $g_{m1}$, $g_{m2}$, and $g_{m3}$ by properly choosing the channel widths of M1 to M6 and the DC bias currents $I_{13}$, $I_{14}$, and $I_{15}$ through SPICE simulation with Level 3 models.
When the output currents of pairs one, two, and three are added as shown in Fig. 19, the total output current is

\[ I_{out} = I_{o1} + I_{o2} + I_{o3} \]  

(II.44)

The reason for using the connection for current addition shown in Fig. 19, rather than that in Refs. [42, 60], is to make the currents in M7 and M8 symmetrical in order to facilitate the design of a differential output structure. Note that the currents in M7 and M8 in Fig. 3 of [42] are not symmetrical, so only a single-ended output OTA can be built.

For SPICE simulation, the MOSFETs used are the same as those used before. The optimum widths of the transistors are shown in Fig. 19 and all channel lengths are 3.5 \( \mu \text{m} \). The power supply is \( \pm 2.5 \text{ V} \), and the bias current \( I_s = 120 \mu \text{A} \). For this choice, the total output current \( I_{out} \) in Fig. 21a indicates that the circuit has almost perfect linearity over a very large differential input range, and the corresponding transconductance \( g_m \) shows that the nonlinearity of \( g_{m1} + g_{m2} \) in Eq. (II.42b) is effectively canceled by \( g_{m3} \).

In Fig. 21a, peaks <1>, <2>, and <3> of the \( g_m \) curve are generated by the centers of the linear ranges of \( I_{o1}, I_{o2}, \) and \( I_{o3} \), respectively. Peak <4> is caused by the overlap of the nonlinear ranges of \( I_{o1} \) and \( I_{o3} \), and peak <5> by those of \( I_{o2} \) and \( I_{o3} \). By this arrangement, the linearity error is controlled to \( \pm 0.35\% \) over an almost \( \pm 3 \text{ V} \) p-p differential input range as is shown by the simulated results in Fig. 21b.

Due to the limitations of the real current sources (M13 to M16) used, the differential linear input range is restricted to \( \pm 2 \cdot (V_{DD} - V_T - V_{ds14,sat}) \). Thus, if the channel width of M16 is large enough to make \( V_{gs16} \) and, therefore, \( V_{gs14} \) close to the threshold voltage \( V_T \), \( V_{ds14,sat} \) can be as low as the minimum value of \( V_{ds,sat} \) which typically equals 0.25V. Therefore, the maximum differential linear input voltage that can be reached for this circuit with \( V_{DD} = -V_{SS} = 2.5 \text{ V} \) is approximately \( 2 \cdot (\pm 1.5) = \pm 3 \text{ V} \). This conclusion is verified by the simulation results in Fig. 21a.
Figure 21. The simulation results for the circuit in Fig. 19 with $V_{DD} = -V_{SS} = 2.5$ V and $I_s = 120 \mu$A. (a) The transfer curve of the total output current $I_{out}$ and the total transconductance $g_m$. (b) Linearity error of $I_{out}$ expressed as a percentage of the full-scale current $I_{14}$. 
Note from Eq. (II.39) that a wide linear input range requires a large value of \( n \) to obtain a large \( V_s \) (\( n \approx 22 \) in this example); but, a large \( n \) may result in large process tolerances which are difficult to control in practice. As larger process tolerances lead to worse linearity, there exists a trade-off between the width of the linear input range (requirement of a large \( n \)) and the sensitivity of linearity (need for a small \( n \)). If it is possible to choose model parameters, one may use a low transconductance parameter \( k \). The reason is that, from Eq. (II.39), a small \( k \) results in larger \( V_s \), which in turn leads to a larger linear input range (but also a lower value of \( g_m \)).

An alternative method leading to small \( n \) is *current subtraction* in which the current of the differential pair three (M5, M6) is subtracted from the sum current of other two mismatched differential pairs instead of added as done in current addition. This scheme can easily be implemented by interchanging the drain terminals of M5 and M6, i.e., by connecting the drain of M5 to the drains of M1 and M3, and the drain of M6 to those of M2 and M4. As was pointed out earlier, for small \( n \) the nonlinear range of one mismatched pair overlaps the linear range of the other mismatched pair, resulting in a large nonlinearity in \( g_{m1} + g_{m2} \) near the center region, \( V_i = 0 \); this situation is shown in Fig. 22 by the curve \( g_{m1} + g_{m2} \) with \( n = 4 \). In that figure, it is also indicated that the nonlinearity of \( g_{m1} + g_{m2} \) can be effectively canceled by subtracting a properly chosen \( g_{m3} \).

Equation (II.39) shows that small \( n \) results in small \( V_s \) and, in turn, a small linear input range. On the other hand, large \( n \) leads to \( g_{m1} + g_{m2} \) having the shape as shown in Fig. 20, only the current addition method can be used and current subtraction method leads to poorer linearity. So, the conclusion is that, if current subtraction method is used, \( n \) cannot be too large and, also, \( n \) cannot be too small. Therefore, there exists an optimum value of \( n \) for the best cancellation of nonlinearities over a maximum linear input range. Note that this situation is different from that in *current addition* where a range of values of \( n \) can be used to achieve a desired input range.
Figure 22. SPICE simulation results for the transconductances $g_{m1}$, $g_{m2}$, $g_{m3}$, and $g_{m1} + g_{m2}$ in the OTA with current subtraction.

It can be observed from Fig. 22 that the best choice for the largest input range with the best possible linearity is to choose $g_{m1}$ and $g_{m2}$ to have their maximum values approximately at the voltages $-V_2$ and $V_2$, defined in Eq. (II.38c), respectively, which means that $-V_{EM} = -V_2$ for $g_{m1}$; thus

$$\frac{dg_{m1}}{dV_i} = \frac{d^2I_o}{dV_i^2} = 0 \quad \text{at} \quad V_i = -V_2 \quad \text{(II.45)}$$

As mentioned above, in current addition, the "shape" of only $g_{m3}$ needs to be adjusted for the best cancellation because $n$ can be chosen at any value; this is not difficult to implement by SPICE simulation. But, in current subtraction, the "shapes" of $g_{m3}$, $g_{m1}$, and $g_{m2}$ must be adjusted for good compensation; consequently, to find the optimum solution for this multidimensional problem by SPICE simulation without a good starting point is more time consuming.
Using the first-order transistor model and solving a seventh-order equation, the solution of (II.45) can be calculated to be $n = 4.2$. After $n$ is found, the equations determining the best "shape" of $g_m$ are the same as those used in current addition. This means Eq. (II.43), with $V_{EM} = V_2$, is also valid in current subtraction for selecting the widths $W_5$, $W_6$, and $W_{13}$. The calculated results, for $n = 4.2$, give $W_5 = W_6 \approx 1.2W_2$ and $W_{13} \approx 0.3W_{14}$. Of course, as mentioned before, these results are obtained from calculations based on a first-order transistor model and, in practice, have to be adjusted slightly when the circuit is simulated by SPICE with a Level 3 model.

The above optimum numbers were checked by SPICE simulations with the same transistor parameters and the same power supplies used in current addition. For three differential pairs and current sources with the widths as: $W_1 = W_4 = 32 \mu m$, $W_2 = W_3 = 8 \mu m$, $W_5 = W_6 = 10 \mu m$, $W_{14} = W_{15} = 400 \mu m$, $W_{13} = 95 \mu m$, and $W_{16} = 200 \mu m$, the simulation results in Fig. 23 show that the circuit has almost perfect linearity with the linearity error less than $\pm 0.2\%$ over the $\pm 2V$ differential input range. Outside that range, the linearity becomes bad very quickly, because the operations of the input transistors are out of the optimum range. Comparing these results with those in Fig. 21a for current addition leads to the conclusion that the linearity of current subtraction is better than that of current addition, but the linear range of the former is less than that of the latter. This result is as expected and can be used as a trade-off in OTA design.

In the simulations, $n = 4$ was used which is very close to the predicted number, and also $W_5$ is close the optimum value. Only $W_{13}$ is less than the optimum number by a factor 0.2; the discrepancy is caused by the first-order transistor model used in the calculations and also by the nonlinearity in the current mirrors, M7 and M9, and M8 and M10.

II.3.3. Other Considerations

A. Tuning. Although three current sources, M13 to M15, are used in both circuits, tuning, as a basic function of OTAs, is not affected and works well. Also, tuning does not
Figure 23. The simulation results for the transfer curve of the total output current $I_{out}$ and the total transconductance $g_m$ of the OTA with current subtraction.

destroy the linearity of the OTAs because the DC bias currents in all current sources can be changed proportionally, via mirroring the master control current $I_s$. The transconductance magnitude can be changed simply by varying the bias current $I_s$ in the current source, or, if this source is built by a transistor $M_s$, by varying the gate voltage of $M_s$.

The simulation results in Figs. 24a and 24b show that the output currents of both circuits can be changed over a wide range by varying $I_s$, and that the linearities of both circuits are maintained over the whole tuning range. Of course, the linear range is reduced with decreasing $I_s$, because reducing $I_s$ decreases the maximum available bias currents of the differential pairs. Note that, in Figs 24a and 24b, $I_{bias}$ represents the DC bias current $I_{14}$ in M14 which is equal to $2I_s$.

The above tuning changes the magnitude of the transconductance and, therefore, varies the frequency parameters of filters, $\omega_c = g_m/C$. As mentioned in the introduction, besides magnitude tuning, the phase of transconductances must be also changeable for
Figure 24. SPICE simulation results for OTA tuning of the magnitude of $g_m$ by varying $I_s$, $I_{bias} = 2I_s$, in Fig. 19. (a) Current addition and (b) current subtraction.
controlling the phase-sensitive quality factors. For this purpose, two RC series branches shown in Fig. 25a are connected between the nodes "a" and "b", and between "c" and "d" of the OTA (see Fig. 19), respectively. In Fig. 2.25b, the resistor R is replaced by a NMOSFET $M_R$ operating in the triode region, so that the equivalent RC time constant, which determines the phase of the transconductance, can be varied by the DC control voltage $V_C$. To be able to verify this conclusion, an integrator is built by the OTA in Fig. 19, terminated in a capacitor, as shown in Fig. 3b, with the value $C_L = 5 \text{ pF}$. Choosing $C = 0.3 \text{ pF}$ and $W_R = 5 \mu m$, the simulation results in Fig. 25c show that the phase of the integrator can be effectively changed from five degree phase lag to 6 degree phase lead as $V_C$ varies from 1 V to 2 V.

**B. Bias circuit.** The bias circuit used here is the same as those used in other op-amps or OTAs. Only one item must be pointed out. In this OTA, to obtain the largest linear input range, as discussed above, $V_{gs16}$ should be designed to be close to the threshold voltage. On the other hand, to build a good current mirror, the voltage $V_{ds23}$ should be chosen to be nearly equal to $V_{ds24} = V_{gs24}$. Thus, in order to be able to select $V_{gs16}$ and $V_{ds23}$ without the restriction that $V_{gs16} + V_{ds23} = 2V_{dd}$, the two diode-connected transistors M21 and M22 are used in the bias circuit.

**C. Frequency Response.** Although three differential pairs are used in this OTA, the frequency response is close to that of the regular differential pair. The reason is that this OTA has the same internal nodes as an OTA built by one differential pair, because the three outputs of differential pairs are connected together. Therefore, the frequency response of this OTA is also dominated by the output nodes.

For lengths chosen as $L = 3.5 \mu m$, the cutoff frequency, $f_{-3dB}$, is around 100 MHz for the whole tuning range. For $L = 2.4 \mu m$, $f_{-3dB} \approx 200 \text{MHz}$ (see Fig. 26). As mentioned in the introduction, although AC compensation techniques can be used to improve the frequency response, the ultimate limitation stems from the relatively low intrinsic cutoff
Figure 25. A RC branch used for OTA tuning of the phase of $g_m$; (a) with passive resistor, and (b) with transistor as a simulated resistor. (c) SPICE simulation results for an integrator built with the OTA in Fig. 19 terminated by a capacitor $C_L = 5 \mu F$ (see Fig. 3b).
D. Output Impedance. As mentioned earlier, the output impedance is also a very important parameter of an OTA. For example, in the design of a transconductance-C filter, the quality factor is determined by the output impedance of the transconductors used. If the output impedance is too small, the quality factor will be very poor. So, the higher the output impedance of an OTA, the better the OTA.

Fortunately, the output impedance of a MOSFET is very high. So, usually, there is no special technique required for output impedance enhancement to an OTA. The simulation results for the OTA in Fig. 19 show that, for the whole tuning range, the low-frequency output impedance magnitude is more than 300 kΩ (see Fig. 27) which is high.
enough for the needs of most design requirements.

![Graph](image)

**Figure. 27.** The simulation results for output impedance of the OTA in Fig. 19 with tuning.

**E. Body Effect, DC Offset, Output Swing, PSR, and CMR.** As mentioned in the introduction, the fully-balanced circuit structure can double output swing and increase PSR (power supply rejection) ratio and CMR (common-mode rejection) ratio. Also, if two outputs of OTA have the exact same paths from input to output (no process tolerance), DC offsets appearing at the both outputs are the same. Then, DC offsets can be cancelled when a differential output current is taken. So, no special care is needed for DC offsets. The same situation applies for body effect which usually is a big problem for single-ended output structures.

Of course, in practice, process tolerances cannot be zero. But, because of the fully-balanced structure used, the impact of process variations can be reduced. Particularly, with some special layout techniques, the process tolerances in differential circuits can be reduced to be very small. Note that this situation does not apply to a single-ended
circuits. A detailed discussion will be given in Chapter VI.

Besides the advantages of fully-balanced structures mentioned above, differential circuits also can improve total harmonic distortion by cancelling second-order harmonics [93, 201, 202].

A problem for fully-balanced structures is stability. In Fig. 19, the transistors M17 to M20 are used for CMF (common-mode feedback) to stabilize the common-mode output signals. Because this issue is very important for differential circuits, an entire chapter (Chapter III) will be used for this discussion.

II.4. COMPARISONS

In Sections II.2 and II.3, four OTAs have been designed with different features. For example, the CMOS cross-coupled OTA has good linearity and high common-mode rejection ratio to arbitrary input signals, but its input range is relatively small. The NMOS cross-coupled OTA has very high linearity over a large input range, but it requires purely differential input signals. The OTA based on current addition has good linearity over a very large input range and accepts any type of input signals without symmetry being required. But, the large width ratio of two input transistors in mismatched differential pairs may cause large process tolerances. The OTA based on current subtraction has small width ratio and very good linearity, but its input range is reduced. These circuits provide the choice of different OTAs for different applications and also verify the previous assertion that "it is impossible to design an OTA to satisfy all requirements."

Table I summarizes the comparisons of the differential OTAs.
### TABLE I
THE SUMMARY OF THE COMPARISON OF THE DIFFERENT OTAS DESIGNED IN THIS CHAPTER

<table>
<thead>
<tr>
<th></th>
<th>Linear</th>
<th>CMR with Symmetrical Inputs</th>
<th>CMR with Unsymm. Inputs</th>
<th>Device Mismatch Problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linearity</td>
<td>Input Range</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CMOS coupled</td>
<td>Good</td>
<td>Small</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>NMOS coupled</td>
<td>Very Good</td>
<td>Large</td>
<td>High</td>
<td>Poor</td>
</tr>
<tr>
<td>Current Add.</td>
<td>Good</td>
<td>Large</td>
<td>High</td>
<td>High</td>
</tr>
<tr>
<td>Current Sub.</td>
<td>Very Good</td>
<td>Small</td>
<td>High</td>
<td>High</td>
</tr>
</tbody>
</table>
CHAPTER III

DESIGN CONSIDERATIONS FOR COMMON-MODE FEEDBACK

In the last chapter, the linearity of OTAs has been discussed in detail. Also, some other design criteria such as tuning, frequency response, body effect and DC offset have been briefly considered. In this chapter, the discussion will focus on the common-mode feedback circuit which is a key sub-circuit of fully-differential circuitry.

As mentioned in the introduction and the last chapter, a fully-balanced circuit has many advantages, such as, reducing body effect and DC offset, improving dynamic output swing and total harmonic distortion, increasing PSRR and CMRR, and minimizing effects of process tolerances. Compared to a single-ended output, the structure of a fully-balanced circuit is only marginally more complex; there are no additional elements required, but only one current mirror should be broken to form two current sources, such that two independent outputs can be formed. This step, however, breaks the feedback loop for common-mode (CM) output signals, because the load which crosses two outputs is only a bridge for differential-mode (DM) outputs, but not for CM outputs. This, therefore, results in an undefined CM output voltage and reduced CMR. The most common and direct method for solving this problem is to add an additional common-mode feedback loop, i.e., a CMF circuit to stabilize the CM output signal and, therefore, to improve CMR. To date many CMF circuits have been reported for use in op-amps, but only a few designs are available for OTAs [127, 102]. In [127], a simple and effective CMF circuit was designed for the OTA. However it requires a high-precision fabrication process to guarantee that the currents in the tail current source and in the load current source are approximately the same. In [102], an OTA was designed by using a totally different prin-
ciple, so the tuning method used there cannot be applied to the OTAs discussed in the last chapter. So, new CMF circuits should be designed.

The CMF circuits may increase both power consumption and chip area. Also, because they form additional loads for the differential-mode signal, the frequency response of the OTA will be degraded. To reduce these effects of CMF circuits, the simplest circuitry with small-size transistors should be used. Another potential problem is caused by the nonlinearity of a CMF circuit. After CMF is added to an amplifier, the output voltage of the amplifier will be constrained by the currents in the CMF circuit. Due to the nonlinearity of the I-V characteristic of MOSFETs, the output voltages may become asymmetric [195]. So, choosing a CMF circuit with minimum nonlinearity distortion is also an important consideration for the design. Furthermore, when a CMF circuit is used in an OTA, a further serious problem may appear: usually, both CMF and tuning implement their functions by changing DC bias currents. Therefore, the operation of CMF may cause undesired tuning, and the operation of tuning may result in a large DC offset. Furthermore, this interdependence of two operations may lead to the filtering system being unstable. Consequently, to make the operations of CMF and tuning independent of each other is a very important aspect for fully differential OTA design.

In Section III.1, three popular CMF circuits used in operational amplifiers (op-amps), will be analyzed and compared in terms of power consumption, chip area, parasitic capacitance and nonlinearity, and one of them will be chosen as a candidate for OTA design. In Section III.2, the improved versions of this candidate will be proposed so that the interdependence between tuning and CMF is minimum.

III.1. BASIC PRINCIPLES

In a fully differential amplifier, the CMF circuit senses the CM output voltage $V_{o,CM}$ but, ideally, not the DM output voltage $V_{o,DM}$. The output signal of the CMF cir-
cuit is fed back to the amplifier to adjust the DC current in the output stage so that $V_{o,CM}$ can be maintained at a preset DC level, normally at zero.

The three most common CMF circuits are shown in Fig. 28, among which Circuit 1 is a typical differential pair. Because the output voltage $V_s$ is taken from the common sources of transistors $M_1$ and $M_2$, $V_s$ ideally does not change with $V_{o,DM}$ of the amplifier, but only with $V_{o,CM}$. In Circuit 2, where the voltage $V_{cm}$ is the preset DC output level, changes of the currents $I_{d1}$ and $I_{d4}$ caused by $V_{o,DM}$ are balanced by corresponding changes of the currents $I_{d2}$ and $I_{d3}$. For $V_{o,CM}$, this balance is broken, and the voltage $V_s$ changes until a new balance is reached [192]. In Circuit 3, transistors $M_4$ and $M_5$ are part of the output stage of the amplifier, and transistor $M_6$ is part of the bias circuit. Transistors $M_1$ and $M_2$ operate in the triode region, and $M_3$ is used to set a reference voltage $V_{cm}$. It can be seen that, in Circuit 3, the sum $I_{d1} + I_{d2}$ is constant with $V_{o,DM}$ changing, because $M_1$ and $M_2$ are operated in the triode region and the current changes in $M_1$ and $M_2$ are in the opposite direction. But the sum of $I_{d1} + I_{d2}$ varies with $V_{o,CM}$, because both currents in $M_1$ and $M_2$ are changed in the same direction.

Understanding the basic principle of CMF circuits, a comparison can be made. For power consumption, it can be seen from Fig. 28 that Circuit 1 consumes little power because the current $2I_{dd}$ can be very small. Circuit 2 consumes more power than Circuit 1 since two differential pairs are used. For Circuit 3, the additional power consumption is almost zero, because $M_1$ and $M_2$ are inserted to the output stage of the amplifier and $M_3$ is inserted to the bias circuitry.

As to chip real estate, Circuit 1 occupies a very small area, because it consists of only three small-size transistors. Circuit 2 contains eight transistors, so that, compared to Circuit 1, the area increases nearly three times. The chip area for Circuit 3 is large, because large currents flow through the output stage and bias circuit and the transistor sizes for $M_1$, $M_2$, and $M_3$ must be large.
Figure 28. The three most common CMF circuits used in op-amps.
In Circuits 1 and 2, as mentioned above, the sizes of input transistors $M_1$ to $M_2$, and $M_1$ to $M_4$, respectively, can be very small, so that the additional parasitic capacitive load caused by the CMF circuits is very small. But in Circuit 3, large transistor sizes of $M_1$ and $M_2$ are used, so that the additional parasitics are large also.

The CMF loop gain is required to be as large as possible because it determines the sensitivity of the circuit to the CM signal. For Circuit 1, the output signal $V_s$ of the CMF circuit is used to control the gate of the transistors which act as the current sources in the output stage of the amplifier (the transistor $M_4$ in Fig. 30) and which must operate in saturation. Therefore, the current $I_s$ in the output stage changes with $V_s$ according to the square-law characteristics of a MOSFET. This situation is the same for Circuit 2, because the output of Circuit 2 is also connected to the gate of current source transistors [192]. Because transistors $M_1$ and $M_2$ in Circuit 3 operate in the triode region, the change of $I_s$ caused by the change of CM output voltage is less than that for Circuit 1 and its loop gain is smaller.

Next, distortion of the output voltage caused by the nonlinearity of the CMF circuit will be analyzed. Suppose that the preset DC voltages at the output nodes are zero. Then, for Circuit 1, when $V_o^+ = V_o^- = 0$

$$I_{d1} = I_{d2} = I_{dd} = k(-V_s)^2$$  \hspace{1cm} (III.1)

where $k$ and $V_T$ were defined earlier and $V_s0$ is the DC value of $V_s$.

Suppose that for a "pure" DM output voltage ($V_o^+ = -V_o^- \neq 0$) $V_s$ has a small change $\Delta V_s$; then $V_{gs1}$ and $V_{gs2}$ are

$$V_{gs1} = V_o^+ - (V_s0 + \Delta V_s) \quad \text{and} \quad V_{gs2} = -V_o^- - (V_s0 + \Delta V_s)$$  \hspace{1cm} (III.2)

According to Eq. (III.2), $I_{d1}$ and $I_{d2}$ become

$$I_{d1} = k[(V_o^+ - \Delta V_s)^2 + (-V_s0 - V_T)^2 + 2(V_o^+ - \Delta V_s)(-V_s0 - V_T)]$$  \hspace{1cm} (III.3a)

and
$I_{d2} = k \left[ (-V_{o+} - \Delta V_s)^2 + (-V_{s0} - V_T)^2 + 2(-V_{o+} - \Delta V_s)(-V_{s0} - V_T) \right]$ (III.3b)

Considering that $I_{d1} + I_{d2} = 2I_{dd}$ from Eq. (III.1), $\Delta V_s$ is solved from Eq. (III.3)

$$\Delta V_s = (V_{s0} - V_T) - \sqrt{(V_{s0} - V_T)^2 - (V_{o+})^2}$$ (III.4)

It is seen from Eq. (III.4) that, for Circuit 1, $\Delta V_s = 0$ when $V_{o+} = V_{o-} = 0$. But, when a "pure" DM output voltage $V_{o+} = -V_{o-} \neq 0$ is present, $\Delta V_s \neq 0$. When this change of $V_s$ is fed back to the amplifier to vary the output voltage, the latter becomes asymmetric. This means that the pure DM output voltage is distorted. In order to determine the degree of the distortion, consider the relationship between $V_{o+}$ and $V_{o-}$ for $\Delta V_s = 0$. The difference between these two voltages will be a measure of the nonlinear distortion to a pure DM output signal.

If $V_{o+} \neq 0$ and $V_{o-} \neq 0$, and $V_s = V_{s0}$ is considered constant ($\Delta V_s = 0$), $V_{gs1}$ and $V_{gs2}$ are

$$V_{gs1} = V_{o+} - V_{s0} \quad \text{and} \quad V_{gs2} = V_{o-} - V_{s0}$$ (III.5)

This yields

$$I_{d1} = kV_{o+}^2 + 2kV_{o+}(-V_{s0} - V_T) + k(-V_{s0} - V_T)^2$$ (III.6a)

and

$$I_{d2} = kV_{o-}^2 + 2kV_{o-}(-V_{s0} - V_T) + k(-V_{s0} - V_T)^2$$ (III.6b)

From Eq. (III.6) with Eq. (III.1), it can be shown that

$$V_{o+}^2 + V_{o-}^2 + 2(V_{o+} + V_{o-})(-V_{s0} - V_T) = 0$$ (III.7)

If $V_{o+} = -V_{o-}$, Eq. (III.7) is no longer satisfied. So, $V_{o+} \neq -V_{o-}$. The conclusion is that if $\Delta V_s = 0$ for no distortion, $V_{o+} \neq -V_{o-}$. Or, if $V_{o+} = -V_{o-}$ for a purely differential output signal, $\Delta V_s \neq 0$ resulting in distortion.

From Eq. (III.7), the following equation can be generated:
In **Eq. (III.8)**, the ratio \( r = \frac{|V_{o+}|}{2(-V_s - V_T)} \) is a measure of the degree of the distortion. If \( r \ll 1 \), \( V_o = -V_o \). This implies that the distortion can be reduced if \( V_s \) is chosen to be much larger than \( V_o \) or \( V_o \). It can also be seen from **Eq. (III.4)** that \( \Delta V_s \approx 0 \) if \( r \ll 1 \).

This conclusion has been verified by SPICE simulations with the transistor parameters as used before. The simulation results in Fig. 29 show that the nonlinear distortion for \( V_s = -3.5 \) \( V \) is less than that for \( V_s = -2.5 \) \( V \).

![Graph showing simulation results for common-mode output voltages](image)

**Figure 29.** SPICE simulation results for common-mode output voltages, \( V_{o+} \) and \( V_{o-} \), with different \( V_s \).

Applying the same analysis to Circuit 2 shows that \( V_o = -V_o \), i.e., there is no distortion. For Circuit 3, distortion appears also. It was pointed out in [195] that SPICE simulation results indicate that Circuit 2 has the best linearity and Circuit 3 has the worst.
The above analysis is summarized in Table II, from which it can be seen that the most important problem of Circuit 2 is the large chip area. Because an OTA-C filter may require many OTAs, the OTA circuit should be designed to be as simple as possible. For Circuit 3, the main drawback is the large parasitics which may degrade the high-frequency response. The only disadvantage of Circuit 1 is potential nonlinearity which, fortunately, is less critical in OTA-C filters, because (i) the output voltage swing of OTAs in an OTA-C filter is generally not large and (ii) \( |V_{o0}| \) can be designed large by using a large-size transistor for the current source \( 2I_{dd} \) and small-size transistors for \( M_1 \) and \( M_2 \) in Circuit 1. Circuit 1 is, therefore, recognized as the best candidate for CMF circuitry in OTAs and will be modified so that it can be used in OTAs.

TABLE II
THE SUMMARY OF THE COMPARISON OF THREE TYPICAL CMF CIRCUITS

<table>
<thead>
<tr>
<th></th>
<th>Add’l Power Consumption</th>
<th>Chip Area</th>
<th>Linearity</th>
<th>Sensitivity to CM signal</th>
<th>Parasitic Capacitance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Circuit 1</td>
<td>Small</td>
<td>Small</td>
<td>Poor*</td>
<td>Large</td>
<td>Small</td>
</tr>
<tr>
<td>Circuit 2</td>
<td>Large</td>
<td>Large</td>
<td>Good</td>
<td>Large</td>
<td>Small</td>
</tr>
<tr>
<td>Circuit 3</td>
<td>Zero</td>
<td>Large</td>
<td>Poor**</td>
<td>Small</td>
<td>Large</td>
</tr>
</tbody>
</table>

* If \( |V_{o0}| \) is chosen to be very large and output swing is small, the linearity is acceptable.

** See Reference [195].

III.2. IMPROVED CMF CIRCUITS

A typical simple fully-differential OTA with the above-chosen CMF circuit is shown in Fig. 30; the DC output voltage \( V_{o+} \) and \( V_{o-} \) can be expressed as
Figure 30. Circuit diagram for a simple fully differential OTA with a typical common-mode feedback circuit.

\[ |V_{o+}| = |V_{o-}| = V_{gs1} + V_{gs4} + V_{ss} \]  

(III.9a)

Because

\[ V_{gsi} = \left( \frac{2L}{kW_i} I_1 \right)^{1/2} + V_{Ti} \]  

(III.9b)

with \( W_i \) and \( L \) being the widths and length of transistors \( M_i \), then,

\[ |V_{o+}| = |V_{o-}| = \left( \frac{2L}{kW_1} I_1 \right)^{1/2} + \left( \frac{2L}{kW_4} I_2 \right)^{1/2} + V_{T1} + V_{T4} + V_{ss} \]  

(III.9c)

In Eq. (III.9c), \( V_{o+} \) (or \( V_{o-} \)) is a monotonic function of bias current \( I_s \). In op-amps, because \( I_s \) is a constant, no DC offset is caused. But in an OTA \( I_s \) must be changeable for post-fabrication tuning, therefore \( I_s \) is not a constant. When \( I_s \) increases, \( V_{gs4} \) increases. This results in \( V_{o+} \) increasing because \( V_{gs1} \) is a constant. So, DC offsets are generated on the both outputs during tuning. In order to avoid or minimize this prob-
lem, the CMF circuit in Fig. 30 must be modified.

Equation (III.9c) points to three possibilities to improve the CMF circuit in Fig. 30. First, $V_{o+}$ is a monotonic function of $I_s$ because $I_1$ remains constant when tuning. If $I_1$ can be made to change as a function of $I_s$, but in direction opposite to the second term, the changes of the second term may be canceled by the change of the first term and $V_{o+}$ may become near constant. Second, because during tuning, $g_m$ changes with the DC currents in the input transistors, it is not required to change the currents, say $I_s$, in the output stage. Thus, if a new tuning method can be designed to change $g_m$ without changing $I_s$ in the output stage, the second term of Eq. (III.9c) becomes constant. This results in a constant $V_{o+}$. Third, if an additional adjustable bypass current from the output node to ground is introduced, $I_s$ will not be controlled by $V_s$, but by a bias voltage. Then, the CMF loop does not include the transistor M4, and, therefore, $I_s$ can be removed from Eq. (III.9c) and $V_{o+}$ becomes constant. Based on these considerations, three improved circuits are proposed in Figs. 31, 33 and 34.

Figure 31 shows the improved CMF circuit for the first choice mentioned above [202]; it adds only one transistor M5 to the CMF circuit in Fig. 30. Because the gates of M5 and M5 are connected, the current $I_5$ changes proportionally to $2I_s$, and $I_1$ becomes

$$I_1 = \frac{(I_3 - I_5)}{2} = \frac{I_3}{2} - \frac{I_s W_5}{W_s}$$

(III.10)

In Eq. (III.10), $I_1$ is a decreasing function of $I_s$ as required, which results in Eq. (III.9c) becoming

$$V_{o+} = \left[ \frac{L}{kW_1} (I_3 - \frac{2W_5 I_s}{W_s}) \right]^{1/2} + \left( \frac{2L}{kW_4 I_s} \right)^{1/2} + V_{T1} + V_{T4} + V_{ss}$$

(III.11)

It is apparent from Eq. (III.11) that, for a proper choice of the width of M5, $W_5$, or, equivalently, $W_3$, i.e. the choice of $I_3$, the increase of the second term may be balanced by the decrease of the first term when $I_s$ increases. Due to the square-root in both
Figure. 3.1. Circuit diagram for the improved CMF circuits suitable for OTAs with tuning; First choice.

the first and the second terms, a complete cancellation is impossible. But, an optimum solution may exist. By setting the derivative of $V_{o+}$ in Eq. (III.11) with respect to $I_s$ to zero shows that $V_{o+}$ has only one extreme value at

$$I_s,m = \frac{W_1 W_2^2 I_3}{2W_4 W_5^2 + 2W_1 W_5 W_s}$$  (III.12)

This means that $V_{o+}$ is either a convex or a concave function of $I_s$. Further, in Eq. (III.11) $V_{o+}$ is monotonic in $I_3$. For a function with these two properties, the range of variation or the extreme maximum (or minimum) value is minimum if $V_{o+}(I_{s,min}) = V_{o+}(I_{s,max})$.

From this condition, the optimum $I_3$ can be solved to be

$$I_3 = \left[ \frac{W_5}{W_s} \left( \frac{W_4}{2W_1} \right)^{1/2} \left( \sqrt{I_{s,max}} + \sqrt{I_{s,min}} \right) + \left( \frac{W_1}{2W_4} \right)^{1/2} \left( \sqrt{I_{s,max}} - \sqrt{I_{s,min}} \right) \right]^2 + \frac{2W_5 I_{s,min}}{W_s}$$  (III.13)

After $I_3$ is obtained, it is easy to calculate the required width $W_3$ for a constant gate vol-
This circuit and the circuit in Fig. 30 were simulated by SPICE to verify the above conclusion. The transistor widths [μm] are shown in Fig. 30 and Fig. 31, and the transistor lengths are 10 μm. Note that, in Fig. 30, \( W_1 = 3 \) μm and \( W_3 = 20 \) μm, which shows \( W_3 \neq 2W_1 \). The reason for this choice is to obtain large \( V_{gs1} \) and \( V_{gs2} \) to decrease nonlinear distortion as discussed in the last section. In Fig. 31, \( W_3 = 28 \) μm is chosen according to \( I_3 = 53.9 \) μA which is calculated by using Eq. (11.13) with \( I_{s,min} = 100 \) μA and \( I_{s,max} = 250 \) μA. With these choices, the simulation results in Fig. 32 show that, \( V_{o+} \) changes from -30 mV to 600 mV (curve \(<1>\)) with \( I_s \) changing for the CMF circuit in Fig. 30, but changes from -30 mV to only 50 mV (curve \(<2>\)) for the circuit in Fig. 31. DC offset is reduced by nearly a factor 10, which is generally acceptable in the design of fully-differential OTAs.

![Figure 32. SPICE simulation results for common-mode output voltage, \( V_{o+} \), of the OTA with the CMF circuits \(<1>\) in Fig. 30 and \(<2>\) in Fig. 31.](image)
The maximum value of $V_{o+}$ at $I_{s,m}$ in Curve <2> cannot be reduced. It can be seen that, after $I_{s,m}$ is substituted in Eq. (III.11), $V_{o+}$ is only function of $W_i$ because $I_3$ has been set for the condition $V_{o+}(I_{s,min}) = V_{o+}(I_{s,max})$. Then, if $W_i$ is changed, not only Curve <2> changes but also Curve <1>. This means the maximum value of $V_{o+}$ in both curves change proportionally, resulting in the ratio of the two $V_{o+,max}$ remaining almost constant.

Observe from Curve <2> in Fig. 32 that $V_{o+}(I_{s,min}) \approx V_{o+}(I_{s,max})$. This means the current $I_3$ as calculated by Eq. (III.13) is close to $I_3$ required in the simulations. Note, though, that Eq. (III.13) should be modified by including the channel-length modulation parameter $\lambda$ if small-length transistors are used.

The other considerations discussed in the last section are still satisfied for the improved circuit in Fig. 31; neither power consumption nor chip area increase significantly because $I_5$ in Fig. 31 can be chosen as small as $I_1$. Also, because the output signal does not flow through $M_5$, the output parasitics of OTA and linearity are not changed.

For the second choice, the corresponding proposed circuit is shown in Fig. 33 [199] where the actual CMF circuit has not changed; there are only two bypass transistors $M_6$ and $M_7$ placed in parallel to the input transistors $M_{d1}$ and $M_{d2}$. Now, as $V_a$ changes, the current $I_6$ changes, which, in turn, results in $I_{d1}$ changing as

$$I_{d1} = I_s - I_6$$  \hspace{1cm} (III.14)

Since the transconductance $g_m$ is proportional to $I_{d1}$, but not $I_s$, then, $g_m$ can be tuned without changing $I_s$, which results in constant $V_{o+}$ by Eq. (III.9c). Because this is obviously true, a proof by SPICE simulation is not necessary. This configuration will be used in the GaAs OTA which will be discussed in Chapter V.

For the last choice mentioned above, the CMF circuit in Fig. 34, compared to the
Figure 33. Circuit diagram for the improved CMF circuits suitable for OTAs with tuning; Second choice.

Figure 34. Circuit diagram for the improved CMF circuits suitable for OTAs with tuning; Third choice.
CMF circuit in Fig. 30, contains two more transistors $M_8$ and $M_9$ which are connected in parallel to transistors $M_4$ and $M_{10}$ in Fig. 34. The output voltage $V_s$ of the CMF circuit is now used to control the gates of $M_8$ and $M_9$, instead of $M_4$ and $M_{10}$. Then, the output voltage is stabilized not by changing the bias current $I_s$, but by adjusting the current $I_9$, which results in $V_{o+}$ in Eq. (III.9c) becoming

$$V_{o+} = \left( \frac{2L}{kW_1} I_1 \right)^{1/2} + \left( \frac{2L}{kW_9} I_9 \right)^{1/2} + V_{T1} + V_{T9} + V_{as}$$  \hspace{1cm} (III.15)

In Eq. (III.15), $V_{o+}$ is independent of $I_s$, so, $V_{o+}$ is, ideally, constant when $I_s$ changes. But, because $I_9$ is not totally independent of $I_s$ due to $I_s + I_9 = I_{11}$, $I_s$ changes will cause DC offset. It should be pointed out that, obviously, this dependence is determined by the ratio $I_9/I_4$, so, an acceptable DC offset can be obtained by a proper choice of $W_4/W_9$. If $W_4/W_9$ is chosen to be very large, $I_9$ is almost independent of $I_s$. Therefore, the CMF loop is nearly independent of $I_s$ and only a very small DC offset is generated by tuning. But, due to small transistors being used for $M_8$ and $M_9$, the gain of CMF circuit is small too, which results in a DC offset caused by common-mode output signals. This implies that this circuit cannot used in OTAs with large CM output swing. On the other hand, if large transistors are used for $M_8$ and $M_9$, a DC offset will be generated by tuning, power consumption will be increased, chip area will be large, and parasitics will be big. So, there is a tradeoff. Fortunately, in filtering applications, the CM output swing is generally not large [201] and this is particularly true for filters operating under the low power supplies such as $\pm2.5$ V. So, the circuit in Fig. 34 can be used with a small $W_8$ and $W_9$, related to $W_4$ and $W_{10}$.

III.3. DISCUSSION

The CMF circuit in Fig. 31 is simple and increases neither power consumption nor chip area significantly when $I_s$ in Fig. 31 is chosen as small as $I_1$. Also, because the
AC signal does not flow through the CMF circuit, the output impedance and the frequency response of the OTA are not changed by the added transistor M5. The disadvantages of this circuit are that DC offset cannot be canceled totally because the operation of tuning and CMF are still functions of $I_s$ and this circuit is a little sensitive to process tolerances.

The advantage of the circuit in Fig. 33 is that tuning and the CMF loop are totally independent of each other, so that no DC offset is generated. The disadvantage of the circuit is that the sizes of transistors M6 and M7, usually, should be chosen to be comparable to that of M1 to maintain a large tuning range; thus power consumption and chip area will increase. Also, the additional transistors increase parasitic capacitances at the nodes A and B in Fig. 33, which results in a reduction of bandwidth.

The CMF circuit in Fig. 34 has also a simple structure and increases small chip area and small power consumption when the sizes of the transistors $M_8$ and $M_9$ in Fig. 34 are chosen to be small. The impact to AC performance of the OTA is small because the small transistors are used. The problem of this circuit is the DC offset cannot be totally canceled and this type of CMF circuit cannot be applied to an ASP circuit with large CM signals.

The above comparison is summarized in Table III.
### TABLE III

THE SUMMARY OF THE COMPARISON OF THE DIFFERENT CMF CIRCUITS

<table>
<thead>
<tr>
<th>Circuit</th>
<th>DC Offset (Tuning)</th>
<th>DC Offset (CM Signal)</th>
<th>Circuit Structure</th>
<th>Add'l Power Consumption</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 31</td>
<td>Smaller</td>
<td>Zero</td>
<td>Very Simple</td>
<td>Very Low</td>
</tr>
<tr>
<td>Fig. 33</td>
<td>Zero</td>
<td>Zero</td>
<td>Moderate</td>
<td>High</td>
</tr>
<tr>
<td>Fig. 34</td>
<td>Small</td>
<td>Small</td>
<td>Simple</td>
<td>Low</td>
</tr>
<tr>
<td>Circuit</td>
<td>Add'l Chip Area</td>
<td>Impact to AC Performance</td>
<td>Process Tolerances</td>
<td></td>
</tr>
<tr>
<td>Fig. 31</td>
<td>Very Small</td>
<td>None</td>
<td>Sensitive</td>
<td></td>
</tr>
<tr>
<td>Fig. 33</td>
<td>Large</td>
<td>Large</td>
<td>Small</td>
<td></td>
</tr>
<tr>
<td>Fig. 34</td>
<td>Small</td>
<td>Small</td>
<td>Small</td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER IV

CMOS CONTINUOUS-TIME INTEGRATED FILTERS

In the last two chapters, the design considerations for OTAs have been discussed and several OTAs have been designed with different features shown in Table I. With these OTAs, a filter, as typical analog signal processing circuit, can be designed in transconductance-C ($g_m$-C or OTA-C) form. For the design of filters, it will also be shown how to design voltage-to-current converters, resistors, gyrators, inductors, and biquads, by using transconductance elements and capacitors.

In Section IV.1, the basic principles of filter designs in OTA-C form will be discussed. In Section IV.2, a third-order LC ladder elliptic filter with 20 MHz bandwidth will be designed in 2 μm CMOS technology. The purpose for choosing an elliptic filter for the design is that the transmission zero of this filter, located at a frequency higher than the cutoff frequency, is very sensitive to the parasitics that are very important parameters for integrated circuit design. In Section IV.3, a 4 MHz sixth-order cascade Bessel filter will be designed in 3.5 μm BiCMOS technology. The reason for selecting a Bessel filter is that the phase of this filter is linear, and this filter is widely used in computer disk drives. Both of these filters are typical for industrial applications and are designed in LC-ladder and cascade configurations.

IV.1. BASIC PRINCIPLES

High-order RLC filters (of order greater than two) are generally built in one of two configurations: cascade structure by cascading simple second-order building blocks or LC ladder structure [12]. The procedures for the design of these types of filters have
been well developed in the last four decades. Fortunately, transconductance-C filters can be designed by using either of these two procedures.

In the design of a high-order cascade filter, a second-order filter, called a biquadratic filter or biquad, is used as a basic building block [1]. Usually, the biquadratic filter is built by two-integrator loops or by a simulated RLC filter. As discussed before, either an integrator or an RLC filter can be built easily by using OTAs and capacitors. So, a cascade filter can be built as an OTA-C structure.

In the design of an LC ladder filter, inductors and capacitors are the basic building blocks. As discussed above, capacitors are very easily built in integrated form, and inductors can be simulated by OTAs and capacitors. Therefore, an LC ladder filter can be simulated by OTAs and capacitors. Because only inductors/resistors are replaced by the simulated inductors/resistors in the design of a simulated LC ladder filter, this method is also called "element substitution." In this section, both methods will be introduced.

IV.1.1. Cascade Filter Design

A filter is a two-port network that shapes the spectrum of the input signal in order to obtain an output signal with the desired frequency content (see Fig. 35a). Generally, the performance of this network is represented by a voltage transfer function, \( H(s) \), as follows:

\[
H(s) = \frac{V_o(s)}{V_i(s)} = \frac{a_ms^m + a_{m-1}s^{m-1} + \cdots + a_0s^0}{b_ns^n + b_{n-1}s^{n-1} + \cdots + b_0s^0}
\]

where \( a_i \) and \( b_i \) are constants, and \( V_o(s) \) and \( V_i(s) \) are the Laplace-transformed terminal voltages of the filtering network, respectively.

In a filter, \( n \) is always greater than \( m \) and is called the order of the filter. In the frequency domain, the transfer function of a filter is, usually, represented by two important functions: \( |H(j\omega)| \) is the magnitude of the filter and \( \phi(\omega) \) is the phase of the filter.
A filtering network can be classified into four different types: lowpass, highpass, bandpass, and bandreject. The function of a lowpass filter is to pass low frequencies from DC to a desired cutoff frequency \( \omega_c \) and to attenuate the frequencies above \( \omega_c \). The function of a highpass filter is to pass high frequencies above \( \omega_c \) and to attenuate the frequencies below \( \omega_c \). The function of bandpass filter is to pass a finite band of frequencies (above \( \omega_{OL} \) and below \( \omega_{OH} \)) and to attenuate both lower and higher frequencies. The function of the bandreject filter is to attenuate a finite band of frequencies and to pass both lower and higher frequencies. The first three filters are the most popular ones; their characteristics are shown in Fig. 36.

A filter of 1st- or 2nd-order is called low-order filter and a filter with higher than 2nd-order called high-order filter. A low-order filter can be built by integrators and a high-order filter usually is designed by cascading several low-order filters.

Mathematically, Eq. (IV.1) can be rewritten as:
Figure 36. The filter characteristics of lowpass (a), highpass (b), and bandpass (c), where the solid lines are for the magnitude and the dashed lines for the phase.

\[ H(s) = \prod_{i=1}^{n/2} H_i(s) = \prod_{i=1}^{n/2} \frac{\alpha_2 s^2 + \alpha_1 s + \alpha_0}{s^2 + s \omega_{oi} Q_{oi} + \omega_{oi}^2} \] for even \( n \) (IV.2a)

\[ H(s) = \frac{k_1}{s + k_2} \prod_{i=1}^{(n-1)/2} H_i(s) = \frac{k_1}{s + k_2} \prod_{i=1}^{(n-1)/2} \frac{\alpha_2 s^2 + \alpha_1 s + \alpha_0}{s^2 + s \omega_{oi} Q_{oi} + \omega_{oi}^2} \] for odd \( n \) (IV.2b)
Figure 36. The filter characteristics of lowpass (a), highpass (b), and bandpass (c), where the solid lines are for the magnitude and the dashed lines for the phase. (Continued)

where $\alpha_{2i}, \alpha_{1i}, \omega_{oi}, \omega_{oi}, k_1, k_2,$ and $Q_{oi}$ are all constants.

Equation (IV.2) shows that the transfer function, $H(s)$, of an n-th order filter can be decomposed into the product of the transfer functions, $H_i(s)$, of first- or second-order sections. With this type of form, Eq. (IV.2) can be easily implemented as the circuit shown in Fig. 35b with (for n even)

$$H(s) = \frac{V_o}{V_i} = \frac{V_o}{V_{n/2-1}} \frac{V_{n/2-1}}{V_{n/2-2}} \cdots \frac{V_2}{V_1} \frac{V_1}{V_i}$$

The circuit in Fig. 35b shows an easy way to build an n-th order filter by second-order filters.

As mentioned above, a second-order section (a biquad) can be built as a two-integrator loop filter (see Fig. 37a) or as a parallel RLC filter (see Fig. 37b). Both of these circuits can be implemented by only OTAs and capacitors (see Fig. 38).
As discussed in the introduction, in order to simplify OTA-C circuit design, OTAs with the same transconductance are usually employed. If using this condition in the circuits in Figs. 38a and 38b, i.e. $g_{mi} = g_m$, all $i$, both circuits have the same bandpass and lowpass transfer functions which can be expressed as follows:

$$H_{bp}(s) = \frac{V_{bp}}{V_i} = \frac{s \frac{2g_m}{C_2}}{s^2 + s \frac{2g_m}{C_2} + \frac{4g_m^2}{C_1C_2}}$$

(IV.4a)
Figure 38. Fully differential OTA-C implementations of the biquadratic filters in Fig. 37a (a) and in Fig. 37b (b).

\[ \frac{H_{lp}(s)}{V_i} = \frac{-\frac{4g_m^2}{C_1 C_2}}{s^2 + \frac{2g_m}{C_2} + \frac{4g_m^2}{C_1 C_2}} \]  \hspace{1cm} (IV.4b)

Compared to the standard forms for second-order filters with \( \omega_o \) as cutoff frequency, \( Q_o \) as quality factor, \( H_o \) as the bandpass filter’s gain, and \( H_l \) as the lowpass filter’s gain.
Another and most important reason for using OTAs with the same transconductance is to reduce the effect of process tolerances. For example, if different transconductances are used in the circuit in Fig. 38b, the transfer functions and filter's parameters become:

\[
H_{bp}(s) = \frac{V_{bp}}{V_i} = \frac{s}{s^2 + s \frac{2g_{m1}}{C_2} + \frac{4g_{m1}g_{m2}}{C_1C_2}} \quad (IV.7a)
\]

and

\[
H_{lp}(s) = \frac{V_{lp}}{V_i} = \frac{-\frac{4g_{m1}g_{m2}}{C_1C_2}}{s^2 + s \frac{2g_{m4}}{C_2} + \frac{4g_{m2}g_{m3}}{C_1C_2}} \quad (IV.7b)
\]
It can be seen from Eq. (IV.8) that the different $g_{mi}$ may cause additional process errors in $Q_o$, $H_o$, and $H_I$.

Comparing the circuit in Fig. 38a to the circuit in Fig. 38b, the latter has a simpler structure than the former but has the same transfer functions. So, in Section IV.3, a simulated RLC biquad will be used for the design of a sixth-order Bessel filter.

**IV.1.2. LC Ladder Filter Design**

One method for the design of LC ladder filters in $g_m$–$C$ form is called element substitution, which means one uses simulated elements to replace passive $R$, $L$, $C$ elements in an LC ladder network. The feature of this method is that all advantages of LC ladder filters are maintained, such as no signal power consumption (lossless), a simple design procedure, and lowest passband sensitivity to element tolerances [1]. The procedure for the design of LC ladder filters is simple and well developed. Using the simulated elements, the topology of the LC ladder filter does not change. So, the procedure can be fully employed for $g_m$–$C$ filter design. A second successful method is the "signal-flow graph" design procedure. In integrated circuits, the element tolerances to process and operating condition variations are very large. In order to reduce the effect of this
tolerance, designing a filter with low sensitivity to tolerances is very important, especially for design automation where there is no special treatment for this purpose. So, an LC ladder is an excellent candidate.

The key point to the design of an LC ladder filter is no signal power consumption, which results in the power delivered to the filter, $P_i$, equal to the power dissipated in the load, $P_L$. Using the notations in Fig. 39, $P_i$ and $P_L$ can be expressed, with $V_s$ as the source voltage, as

$$ P_i = |I_i(j\omega)|^2 \Re\{Z_i(j\omega)\} = \frac{|V_s|^2}{|R_s + Z_i(j\omega)|} \Re\{Z_i(j\omega)\} \quad \text{(IV.9a)} $$

$$ P_L = \frac{|V_o|^2}{R_L} \quad \text{(IV.9b)} $$

It is well known that, the maximum available power at the input of network, $P_{\text{max}}$, occurs when $R_s$ is equal to the input impedance of network $Z_i$. Then,

$$ P_{\text{max}} = \frac{|V_s|^2}{4R_s} \quad \text{(IV.9c)} $$

In LC ladder filter design, the transfer function, $H(s)$, is usually defined as the square root of power transfer ratio, [1]:

$$ H(s) = \left(\frac{P_L}{P_{\text{max}}}\right)^{1/2} = \left(\frac{4R_s}{R_L}\right)^{1/2} \frac{V_o}{V_s} \quad \text{(IV.10)} $$
For a given \( H(s) \), the reflection coefficient, \( \rho(s) \) at the input of the LC filter can be calculated as \[ |\rho(j\omega)|^2 = 1 + |H_p(j\omega)|^2 \] (IV.11)

Next, \( Z_i(s) \) can be expressed as a function of \( \rho(s) \),

\[ Z_i(s) = R_s \left( \frac{1 + \rho(s)}{1 - \rho(s)} \right)^{\pm 1} \] (IV.12)

Mathematically, \( Z_i(s) \) has the same format as that of \( H(s) \). So, \( Z_i(s) \) (or \( 1/Z_i(s) \)) can be decomposed into

\[ Z_i(s) = F_1(s) + \frac{1}{F_2(s) + \frac{1}{F_3(s) + \frac{1}{F_4(s) + \cdots}}} \] (IV.13a)

where

\[ F_i(s) = k_{i0} + k_{i1}s \] (IV.13b)

with \( k_{i0} \) and \( k_{i1} \) real constants corresponding to resistor (conductor) and inductor (capacitor), respectively.

The format of Eq. (IV.13a) is very easy to implement in an LC ladder network and the corresponding circuit is shown in Fig. 40, where \( F_m = Y_m, F_n = Z_n \) with \( m = 1,3,5 \cdots \) and \( n = 2,4,6 \cdots \). For instance, the circuit with the input impedance as

\[ Z_i(s) = \frac{0.375}{s} + \frac{1}{\left( \frac{4.5714}{s} + \frac{1}{\frac{0.5568}{s} + \frac{1}{\frac{46.095}{s} + \frac{1}{\frac{0.06818}{s} \cdots}}} \right)} \] (IV.14)

can be implemented in the LC ladder network shown in Fig. 41 [1].

The above description shows a design procedure for an LC ladder filter; this is to
convert a transfer function into an input impedance function, to decompose the input impedance into the format in Eq. (IV.13), and to implement this equation in the circuit in Fig. 40. The next task is how to simulate the passive elements in Fig. 40 by transconductances and capacitances.

As mentioned in Section II.1, all passive elements can be simulated by transconductors and capacitors. Fig. 42a shows a voltage to current converter, built by a transconductor, with the transfer function

\[ Y_m(s) = \frac{I_o}{V_i} = g_m(s) \]  

\hspace{1cm} (IV.15a)

Figure 42b shows a simulated resistor with the resistance

\[ R = \frac{V_i}{I_i} = \frac{1}{g_m} \]  

\hspace{1cm} (IV.15b)
Figure 42c shows a unity-gain buffer with the transfer function

\[ H(s) = \frac{V_o}{V_i} = 1 \]  \hspace{1cm} (IV.15c)

Figure 42d shows a gyrator-C inductor with the inductance

\[ L = \frac{C}{g_m} \]  \hspace{1cm} (IV.15d)

With these elements, an LC ladder filter can be simulated.

![Diagram](image)

**Figure 42.** Fully differential OTA-C implementations of the basic building blocks in ASP circuits. A voltage-to-current converter (a), a resistor (b), a unity-gain buffer (c), and a gyrator-C inductor (d).

IV.2. A 20 MHz THIRD-ORDER ELLIPTIC CMOS LC LADDER FILTER

As mentioned in the last section, LC ladder filters have the lowest passband sensitivities to element tolerances. Also, among all types of filters, elliptic filters require the
lowest order to implement given magnitude performance specifications; this, therefore, results in simple circuitry. So, in this section, a third-order elliptic low-pass LC ladder filter will be designed to test the effect of parasitics on quality factors and notch depth and, of course, to test the performance of one OTA designed earlier.

One potential problem for the OTA based on current addition in Section II.3 may be the process tolerance, because mismatched differential pairs are used. It is well known that a transistor with a large width has larger processing errors than one with a small width except that some special treatments are involved. Therefore, the process errors of two input transistors of a mismatched differential pair are different even if these two transistors are placed very close in the layout. This results in processing errors not canceling each other as they do in a matched differential pair, and leads to mismatched output currents. But, a careful check of the OTA in Fig. 19 and of the analysis discussed in Section II.3 shows that the total output current $I_o$ of the OTA is actually not proportional to the width ratio $n$ of a mismatched differential pair, but rather to the term $n/(n+1)$. When $n >> 1$, $I_o$ is independent of $n$ and the effect of process tolerances is, in principle, the same as that of an OTA built with matched differential pairs. With this thought in mind and in order to prove this assertion, the OTA in Fig. 19 is selected as the building block of the filter and is chosen for fabrication.

IV.2.1. Filter Design

Equation (IV.16) shows the transfer function of a third-order lowpass elliptic filter with 23 dB minimum attenuation [1]:

$$H(s) = \frac{0.2816(s^2 + 3.2236)}{(s + 0.7732)(s^2 + 0.4916s + 1.1742)}$$  \hspace{1cm} (IV.16)

By using Eq. (IV.11), the reflection coefficient can be found:

$$|\rho(j\omega)|^2 = \frac{\omega^6 - 1.588 \omega^4 + 0.63 \omega^2}{\omega^6 - 1.5087 \omega^4 + 0.1189 \omega^2 + 0.8243}$$  \hspace{1cm} (IV.17a)
This results in

\[ p(s) = \frac{s^3 + 0.794s}{(s + 0.773)(s^2 + 0.4916s + 1.1742)} \quad \text{(IV.17b)} \]

Then, from Eq. (IV.12), the input impedance, with \( R_s = 1 \), can be calculated as follows:

\[ Z_i(s) = \frac{0.6326s^2 + 0.3802s + 0.4541}{s^3 + 0.6326s^2 + 0.5872s + 0.4541} \quad \text{(IV.18)} \]

Equation (IV.18) can be factored to the following format:

\[ Z_i(s) = \frac{1}{1.293s + \frac{1}{\frac{0.8373s}{1 + 0.3102s^2} + \frac{1}{1.293s + 1}}} \quad \text{(IV.19)} \]

Equation (IV.19) can be implemented by an LC ladder filter with a resistive load \( R_L = 1 \). Then, the first term \( 1.293s \) represents a capacitor capacitance \( \hat{C} = 1.293 \), the term \( 0.8373s/(1 + 0.3102s^2) \) represents an inductor \( \hat{L} = 0.8373 \) connected in parallel with a capacitor \( \hat{C} = 0.3705 \), and the last term \( 1/(1.293s + 1) \) represents a capacitor \( \hat{C} = 1.293 \) connected with the load. With this description, the circuit with passive elements \( (R,L,C) \) for the filter with the transfer function in Eq. (IV.16) is shown in Fig. 43.

![Figure 43. A third-order elliptic low-pass filter with passive elements.](image)

To obtain the equivalent circuit in OTA-C implementation for the circuit in Fig. 43, both element substitution and signal-flow graph methods can be shown to result in the
same circuit topology and thus in identical performance. Using for simplicity the element substitution method, Fig. 44 [205] shows the circuit diagram of the designed filter built with OTAs and capacitors. The OTA \( g_{m1} \) is a voltage-to-current converter, the OTAs \( g_{m2} \) and \( g_{m7} \) simulate the source and load resistors in Fig. 43. The capacitor \( C_L \) and the OTAs \( g_{m3}, g_{m4}, g_{m5}, \) and \( g_{m6} \) form a simulated floating inductor. Transistors \( q_{s1} \) to \( q_{s4} \) are used as an output buffer. Note that, because the interesting output signal is a voltage in this filter, source followers can be used as output buffers. By using the same transconductances, \( g_m \), for all OTAs for the reasons discussed before, the transfer function of this filter can be written as:

\[
H(s) = \frac{2g_m C_2}{C_1(C_1 + 2C_2)} \cdot \frac{s^2 + \frac{4g_m^2}{C_2 C_L}}{(s + \frac{2g_m}{C_1})(s^2 + \frac{s}{2g_m C_1(C_1 + 2C_2)} + \frac{8g_m^2}{C_L(C_1 + 2C_2)})}
\]

(IV.20)

Note that, the DC gain in Eq. (IV.20) is one half instead of unity. So, if a unity gain is of interest, the first OTA, \( g_{m1} \), of value \( 2g_m \) should be used.

After getting the circuit, the next task is to calculate the capacitances according to the desired cutoff frequency of the filter and the transconductances of the OTA. But, before starting this calculation, it is necessary to discuss the effects of parasitics on the performance of the filter because the operating frequency considered here is relatively high.

**IV.2.2. Design Considerations for Frequency Responses**

The elements in Fig. 43 are normalized. The denormalized capacitors \( C_i \) and \( C_L \) in Fig. 44 can be obtained from \( \hat{C_i} \) and \( \hat{L} \) in Fig. 43 as

\[
C_i = \frac{2g_m \hat{C_i}}{\omega_o}, \quad i = 1, 2, 3 \quad \text{and} \quad C_L = \frac{2g_m \hat{L}}{\omega_o}
\]

(IV.21)

where \( \omega_o \) is the cutoff frequency of the filter. Because the filter in Fig. 43 is single-ended
Figure 44. Circuit diagram of the filter in Fig. 43 with OTAs and capacitors.
and that in Fig. 44 is a fully balanced circuit, the relationship between $R_S$ and $g_m$ is $R_S = 1/(2g_m)$, which results in the coefficient 2 in Eq. (IV.21).

Further, if a parasitic capacitance $C_{ip}$ is considered to be connected in parallel to $C_i$ at each node, Eq. (IV.21) must be modified as

$$C_i = \frac{2g_m\hat{C}_i}{\omega_o} - C_{ip}, \quad i=1,2,3 \quad \text{and} \quad C_L = \frac{2g_m\hat{L}}{\omega_o} - C_{ip} \quad (IV.22)$$

From $C_i$ in Eq. (IV.22), the cutoff frequency $f_o$ can be written as

$$f_o = \frac{g_m\hat{C}_i}{\pi(C_i + C_{ip})} \quad (IV.23)$$

It is seen from Eq. (IV.23) that the cutoff frequency is determined by the transconductance of the OTAs and by the values of capacitors and parasitics. The smaller the capacitance $C_i$, the higher the cutoff frequency. But, if $C_i$ is chosen too small, the process tolerances of parasitic capacitances $C_{ip}$ become nonnegligible so that the overall performance of the filter is difficult to predict because the parasitics caused by process tolerance are hard to control. Typically, the values of the capacitors $C_i$ should be chosen larger than $C_{ip}$ by a factor, say, four to five; evidently, this sets a limit on the cutoff frequency $f_o$.

In order to increase the obtainable operating frequency, OTAs with small input and output parasitics are required. Also, in order to reduce the parasitics, as few CMF circuits as possible should be used. As mentioned before, a CMF circuit in an OTA introduces an additional load to this OTA, which results in increasing parasitic output capacitance. But, in a filter, often several OTAs can share a CMF circuit if their output nodes are connected together and, therefore, the effect of CMF circuits can be reduced. For example, for the filter in Fig. 44, a CMF circuit can be shared by $g_{m1}$, $g_{m2}$ and $g_{m3}$, one for $g_{m4}$ and $g_{m5}$ and one for $g_{m6}$ and $g_{m7}$. Unfortunately, this method is not very helpful for the filter designed here, because the transistor sizes used in the CMF circuit are very
small. But, it is mentioned here as a general consideration of OTA-C filter design.

The last choice to reduce the total parasitic capacitance is to use passive resistors instead of the OTAs, $g_{m2}$ and $g_{m7}$, because a real resistor, presently, is easily implemented in an IC process with a reasonable area and a small parasitic capacitance. But, after resistors are used, the transfer function in Eq. (IV.20), assuming $R_S = R_L$ and $C_1 = C_3$, becomes

$$H(s) = \frac{2}{C_1(C_1 + 2C_2)} \frac{s^2 + \frac{4 g_m^2}{C_2 C_L}}{(s + \frac{1}{R_S C_1})(s^2 + s \frac{1}{R_S(C_1 + 2C_2)} + \frac{8 g_m^2}{C_L(C_1 + 2C_2)})}$$ (IV.24)

From Eq. (IV.24), it can be seen that three serious problems appear. a) Because $H(0) = g_m R_S$, the magnitude of the transfer function at low frequencies will change with tuning, i.e., with changing $g_m$. b) Eq. (IV.24) shows that the zeros of $H(s)$ are proportional to $g_m$, one pole is independent of $g_m$ and the other poles have a complicated relationship to $g_m$. So, when $g_m$ changes, the poles and zeros will not change proportionally, resulting in a change of the shape of $H(s)$. c) Because the pole quality factor, $Q_2$, equals to $2R_S g_m \sqrt{2(C_1 + 2C_2)/C_L}$, $Q_2$ will not be constant when tuning. Due to these reasons, passive resistors should not be used.

It is seen from Eq. (IV.23) that the cutoff frequency $f_o$ can be increased by increasing $g_m$. This approach is generally limited because to increase $g_m$ in an OTA requires the use of transistors with larger sizes which in turn results in increased parasitic capacitances $C_{ip}$. Then, the capacitors $C_i$ must be increased to guarantee that the effects of process tolerances can be reduced. Therefore, increases in both $g_m$ and $C_i$ do not generally result in a significant change in cutoff frequency $f_o$.

In conclusion, the obtainable frequency range is limited ultimately by parasitic capacitances. Then, an optimum result for $f_o$ can be obtained only by choosing the circuit
capacitors as small as possible, but larger than the total parasitics by about a factor four or five.

It should be pointed out that the above design considerations do not apply only to the filter used as an example above, but to most filters. They form a general conclusion.

IV.2.3. Simulation Results

From Eq. (IV.22), \( C_i \) in Fig. 44 can be calculated as follows:

\[
C_1 = 2\left( \frac{g_m \hat{C}_1}{\omega_o} - 2C_{in} - 3C_{out} \right)
\]

\[
C_2 = \frac{2g_m \hat{C}_2}{\omega_o}
\]

\[
C_3 = 2\left( \frac{g_m \hat{C}_3}{\omega_o} - 2C_{in} - 2C_{out} - 2C_{q1} \right)
\]

\[
C_L = 2\left( \frac{g_m \hat{L}}{\omega_o} - 2C_{in} - 2C_{out} \right)
\]

where \( C_{in} \) and \( C_{out} \) are the differential input and differential output capacitances of OTAs, and \( C_{q1} \) is the gate capacitance of the output buffer.

Before calculating \( C_i \), the values of \( C_{in} \) and \( C_{out} \) should be determined. An easy and accurate way to obtain the values of \( C_{in} \) and \( C_{out} \) is to use SPICE simulation results. Because a different vendor (MOSIS) will be used for fabrication, the new process file results in the different device parameters, compared to those used in Section II.2.1 for simulations (Allegro Microsystems). The new parameters are: \( V_{Tn} = 0.813 \, V \), \( V_{Tp} = -0.970 \, V \), \( k_n = 53.21 \, \mu A/V^2 \), and \( k_p = 19.96 \, \mu A/V^2 \) (see Chapter VI for details).

The new process is a 2 \( \mu m \) technology. As discussed before, an OTA built with short-channel devices has reduced output impedance. In order to improve this, a cascode output stage is used. Other parts of the OTA are exactly the same as the OTA in Fig. 19.

The circuit diagram for the OTA used here is shown in Fig. 45. With the new device parameters and the channel length being 2 \( \mu m \), the new widths for the OTA are:
Circuit diagram of CMOS linear OTA with cascode output
\( W_1 = W_6 = 80 \mu m, \quad W_2 = W_3 = W_4 = W_5 = 3 \mu m, \quad W_7 = W_8 = W_9 = W_{10} = 200 \mu m, \)
\( W_{11} = W_{12} = 200 \mu m, \quad W_{13} = W_{14} = W_{15} = W_{16} = 50 \mu m, \quad W_{17} = W_{18} = 5 \mu m, \)
\( W_{19} = 46 \mu m, \quad W_{20} = W_{22} = 400 \mu m, \quad W_{21} = 70 \mu m, \quad W_{23} = W_{24} = W_{25} = 60 \mu m, \)
\( W_{26} = W_{27} = 200 \mu m, \) and \( W_{28} = 43 \mu m. \) The other conditions are \( I_s = 235 \mu A \) (or the bias current in mismatched differential pair to be 470 \( \mu A \)), \( V_{DD} = -V_{SS} = 5 \) V, \( V_1 = 0 \) V, and \( V_2 = -3 \) V. The simulation results in Fig. 46a show that the input impedance varies with frequency; thus, from a chosen point, say \( r_1, \) \( C_{in} \) can be calculated as:

\[
C_{in} = \frac{1}{2 \pi f |Z_{in}|} \approx 35 \text{ fF} \quad \text{(IV.26a)}
\]

From Fig. 46b, \( C_{out} \) can be calculated as:

\[
C_{out} = \frac{1}{2 \pi f |Z_{out}|} \approx 140 \text{ fF} \quad \text{(IV.26b)}
\]

Besides the capacitances, the performance of this OTA was also checked due to the new process used. Fig. 47 shows that the linearity of this OTA with \( g_m \approx 140 \mu S \) can be adjusted to be as good as what was obtained before (see Fig. 21a). Fig. 46b also shows that, with 2 \( \mu m \) channel length, the low-frequency output impedance of the cascode stage is 450 k\( \Omega \).

With the above values of \( C_{in} \) and \( C_{out} \), and \( g_m \approx 140 \mu S \), the obtainable cutoff frequency is near 20 MHz with the condition discussed above. The final values of \( C_i \) are \( C_1 = 1.90 \text{ pF}, C_2 = 1.65 \text{ pF}, C_3 = 2.08 \text{ pF}, \) and \( C_L = 1.17 \text{ pF}. \)

The filter has been simulated in SPICE to prove the design. The device widths for the output buffer of the filter are \( W_{sq1} = W_{sq2} = W_{sq3} = W_{sq4} = 300 \mu m, \) and the load resistance is 10 k\( \Omega \). The simulation results in Fig. 48a show that this filter has a 20 MHz cutoff frequency response (see Fig. 48b for a passband detail), 0.6 dB passband ripple, and 23 dB stopband attenuation. When \( I_s \) sweeps from 170 \( \mu A \) to 300 \( \mu A \), the cutoff frequency changes from 17 MHz to 23 MHz. It should be pointed out that the DC gain of this filter is near -6.8 dB, because the DC gain is one half and the gain of the source fol-
Figure 46. Input impedance (a) and output impedance (b) of the OTA in Fig. 45. Zod stands for the differential-mode output impedance and Zod for the common-mode output.
Figure 47. Total output current and transconductance of the OTA in Fig. 45.

Figure 48. The transfer function of the third-order lowpass elliptic filter with tuning (a) and passband detail (b).
lower is 0.91.

In Chapter VI, the layout of this filter will be presented. The experimental performance of the filter fabricated by a MOSIS analog process will be shown and be compared with the simulation results.

IV.3. A 4 MHz SIXTH-ORDER BESSEL CMOS FILTER

In the previous section, a 20 MHz LC ladder filter was designed and good simulation results were obtained. In this section, the design of another class of filter, a cascade filter, will be presented.

In the design of this filter, the OTA with current subtraction will be used. There are several reasons for this. First, as stated before, the purpose of designing a filter is not only to show that a high-frequency filter is easy to design in OTA-C form, but also to test
the performance of the OTA itself. So, in this section, we prefer to test a new OTA other than the one used in the last section. Second, the OTA based on the multiple differential pairs has one more advantage which was not mentioned before: the OTA can be used in a low power supply system. This arises because the multiple differential pairs are connected in parallel to each other, not in series, so, the minimum required power supply for this kind of OTA is the same as that for a simple differential pair. Currently, to design a system with a low power supply is a very interest topic, and will be more important as time goes on. So, in order to test this advantage, the filter designed in this section will be for a ±2.5 V power supply, instead of ±5 V. Third, the key interest of a Bessel filter is a flat delay $\tau (\tau = -d\phi/d\omega)$ which requires, ideally, linear phase and zero excess phase shifts. In order to minimize the excess phase shifts, there are two requirements for the OTAs. First, the OTAs used should have very good linearity so that there is no frequency variation in the cutoff frequency. Second, the cutoff frequency of the OTA used should be much larger than the frequencies of interest to the flat delay of the filter. In this design, both requirements are considered; the OTA with current subtraction is used because it has very good linearity.

The process used for the filter designed here is the same as one discussed in Chapter II (Allegro Microsystems). The OTA circuit is shown in Fig. 19 except that the drain terminal of M5 is connected to node "d" and the drain of M6 to node "b" (see Chapter II for the detailed discussion). The transistor sizes are shown in Fig. 19, except that $W_1 = W_4 = 32 \mu m$, $W_2 = W_3 = 8 \mu m$, $W_5 = W_6 = 10 \mu m$, and $W_{13} = 95 \mu m$. Note that $L = 3.5 \mu m$ as used in Chapter II. The simulation results for linearity and tuning are shown in Figs. 23 and 24b.

IV.3.1. Filter Design

The transfer function of a sixth-order Bessel filter can be found from Table III-4 in Appendix III in [1] and is as follows:
\[ H(s) = \frac{10395}{(s^2 + 8.497s + 18.802)(s^2 + 7.471s + 20.852)(s^2 + 5.032s + 26.514)} \]  

Equation (IV.27a) can be decomposed into the following format:

\[ H(s) = H_1(s) H_2(s) H_3(s) \]  

with

\[ H_1(s) = \frac{18.802}{s^2 + 8.497s + 18.802} \]  

\[ H_2(s) = \frac{20.852}{s^2 + 7.471s + 20.852} \]  

\[ H_3(s) = \frac{26.514}{s^2 + 5.032s + 26.514} \]

In the frequency domain, Eq. (IV.27b) can be written as:

\[ H(j\omega) = |H_1(j\omega)||H_2(j\omega)||H_3(j\omega)|e^{-j(\phi_1(\omega) + \phi_2(\omega) + \phi_3(\omega))} \]  

Equation (IV.27b) indicates that this filter can be designed by using the cascade connection shown in Fig. 35b, and each block can be designed as a second-order filter (biquad). As was mentioned in Section IV.1, a parallel RLC biquad has a simpler circuit structure than a two-integrator loop biquad. So, the circuit diagram of the biquad used in this filter is shown in Fig. 38b, which results in the circuit diagram of the complete filter in Fig. 49 [134].

In Fig. 49, the first four OTAs counted from the left are used to build the first biquad, the second four OTAs are for the second biquad, and the third four are for the last biquad. The last block element is not an OTA, but the output buffer which is shown in Fig. 50. From there, it can be seen that the output stage consists of two source followers followed by two open-emitter NPN bipolar transistors. The reason for the bipolar devices is to increase the output drive.
Figure 49: Circuit structure of sixth-order cascade fully differential OTA-C Bessel filter.
As mentioned earlier, tuning is a very important feature for OTA-C systems. In order to change the transconductances of all OTAs in Fig. 49 at the same time, additional 13 NMOS transistors are used to form a one-to-twelve (master to slave) current source. When the current in the master transistor changes, the currents in all slave transistors will change the same amount (see Fig. 49). Note that the biquad in Fig. 38b has two outputs for lowpass and bandpass. Because the Bessel filter designed here is lowpass, the connection between the biquads should be from $V_{lp}$.

It can be found, from Fig. 1-29a in [1], that for a 6th-order Bessel filter with less than 2 dB magnitude attenuation (a conservative consideration for 3 dB attenuation), the normalized frequency $\omega_n$ is 2.2. Then, if the required delay at DC, $\tau_0$, is 85 nS, the normalizing frequency, $\Omega_0$, must be chosen as
Then, the bandedge frequency of the filter is \( \omega_o = \omega_n \Omega_0 = 25.882 \text{ mrad/s} \), which results in \( f_o = 4.12 \text{ MHz} \). Another parameter for a Bessel filter is delay error which, from Fig. 1-29b in [1], is 0.1% for a 6th-order filter and \( \omega_n = 2.2 \).

After obtaining \( \Omega_0 \), the cutoff frequency and quality factor for the second-order filters with the transfer functions in Eq. (IV.28) can be calculated by comparing Eq. (IV.28) to the standard form of a lowpass filter in Eq. (IV.5b),

\[
\omega_{o1} = \sqrt{18.802} \Omega_0 = 51.018 \text{ mrad/s} \quad \text{or} \quad f_{o1} = 8.12 \text{ MHz}
\]

\[
f_{o2} = 8.55 \text{ MHz}
\]

\[
f_{o3} = 9.64 \text{ MHz}
\]

and

\[
Q_{o1} = \frac{\omega_{o1}}{8.497} = \frac{\sqrt{18.802}}{8.497} = 0.510
\]

\[
Q_{o2} = 0.611
\]

\[
Q_{o3} = 1.023
\]

### IV.3.2. Design Considerations and Simulation Results

In the previous section, the filter was designed with a tuning circuit and output buffer. The next task is to calculate the circuit capacitances for SPICE simulations.

Comparing Eq. (IV.28) to Eq. (IV.4), the circuit capacitances for the cell one (C11, C21), the cell two (C12, C22), and the cell three (C13, C23) can be calculated by

\[
\frac{2 \ g_m}{C_{21}} = \frac{8.4967 \ \Omega_0}{8.497} \quad \text{or} \quad C_{21} = \frac{2 \ g_m}{8.4967 \ \Omega_0}
\]

which results with \( g_m = 70 \ \mu \text{S} \) in,

\[
C_{21} = 1.4011 \ \text{pF} \quad C_{22} = 1.59 \ \text{pF} \quad C_{23} = 2.36 \ \text{pF}
\]
and

\[ \frac{4 g_m^2}{C_{11} C_{21}} = \frac{2g_m}{C_{11}} \frac{8.4967 \Omega_0}{18.802 \Omega_0^2} \]  

which results in

\[ C_{11} = 5.38 \text{ pF} \quad C_{12} = 4.266 \text{ pF} \quad C_{13} = 2.26 \text{ pF} \]  

The capacitors obtained above are the actual filter capacitors, i.e., the real circuit capacitors plus the parasitic capacitors. So, a capacitor should be predistorted for parasitics as was done in Section IV.2.

From Fig. 38b, it can be seen that \( C_1 \) is connected in parallel with one input and one output capacitance of the OTA. Because \( V_{ip} \) is the output voltage of the second-order filter, another input capacitance from the next stage is added to the output node too. So, there are in total two input capacitances and one output capacitance connected with \( C_1 \). Similarly, it can be counted that there are two input and three output capacitances connected in parallel with \( C_2 \).

For this OTA, the input capacitance is 20 \text{ fF}, the output resistance is 300 \text{ k\Omega} and the output capacitance is 100 \text{ fF}. Then, it can be calculated that the real circuit capacitors are:

\[ C_{11,r} = 2\left(\frac{C_{11}}{2} - 2C_{in} - C_{out}\right) = 5.10 \text{ pF} \]

\[ C_{12,r} = 3.99 \text{ pF} \quad \text{and} \quad C_{13,r} = 1.98 \text{ pF} \]  

and

\[ C_{21,r} = 2\left(\frac{C_{21}}{2} - 2C_{in} - 3C_{out}\right) = 0.72 \text{ pF} \]

\[ C_{22,r} = 0.91 \text{ pF} \quad \text{and} \quad C_{23,r} = 1.68 \text{ pF} \]
Before simulating this filter with the above capacitors, several design considerations should be analyzed. First, the result of primary interest in this design is a flat delay $\tau$ which, as a function of the phase $\phi$, is very sensitive to parasitics and unavoidable excess phase shifts of the biquads. But, Eq. (IV.4b) which was used to obtain the above capacitors is only a simplified equation. If the output conductance, $g_0$, and the output capacitance, $C_o$, of the OTAs are considered, Eq. (IV.4b) is modified to:

$$H_{lp}(s) = \frac{4 g_m^2}{(C_1 + C_o)(C_2 + 3C_o)} s^2 + 2s \left( \frac{g_m}{C_2 + 3C_o} + \frac{g_0}{C_1 + C_o} + \frac{3g_0}{C_2 + 3C_o} \right) + 4 \left( \frac{g_m^2 + g_0g_m + 3g_0^2}{(C_1 + C_o)(C_2 + 3C_o)} \right)$$

(IV.36)

Generally, $g_m$ is much larger than $g_0$, and the effect of $g_0$ can normally be ignored. But, because the delay error of a Bessel filter is usually required to be very small, say 0.1% for this design, the effect of $g_0$ cannot be ignored. For example, in this design, $g_m = 70 \mu S$ and $g_0 = 1/300 k\Omega = 3.3 \mu S$, so the effect of $g_0$ can be, generally, ignored. But, considering the second term in the denominator of Eq. (IV.36) which affects the quality factor $Q_0$, and in turn, the phase $\phi$, the effect of $g_0$ cannot be ignored because $3g_0$ is 14.3% of $g_m$. So, in practice, the calculated real circuit capacitor $C_{ijr}$ in Eq. (IV.35) should be adjusted to include the effect of $g_0$.

Second, the parasitic output impedance (resistance and capacitance) is not a constant which is used in the above analysis, but is a function of DC bias, load, etc. So, in practice, to calculate the real circuit capacitors using Eq. (IV.36) is very difficult and not accurate. An easy method to design this filter with the consideration of the parasitics is to use the calculated real circuit capacitors (in Eq. (IV.35)) as a starting point for SPICE simulations. An optimal result can be obtained by changing these capacitors until an acceptable delay error is reached.

Third, it can be seen from Eq. (IV.29) that the total phase of the sixth-order filter
is equal to the sum of the phases of each biquad, so the linearity of the total phase of the sixth-order filter is determined by the linearity of the phase of each second-order biquad.

From Eq. (IV.5b), the phase of a second-order biquad can be calculated as

$$\phi(j \omega) = \arctan \left( \frac{\omega \omega_{oi}}{\omega_{oi}^2 - \omega^2 Q_{oi}} \right)$$  \hspace{1cm} (IV.37)

Equation (IV.37) shows that the phase of a biquad is a function of the ratio of the square of the frequencies and is an inverse function of the quality factor $Q_{oi}$; thus the phase is more sensitive to $Q_{oi}$ and less sensitive to cutoff frequency $\omega_{oi}$. So, to maintain the required $Q_{oi}$ it is very important to obtain good linearity of the total phase of the sixth-order filter, but not for the cutoff frequency. On the other hand, if a large cutoff frequency is used so that $\omega_{oi}^2 \gg \omega^2$ is valid, then Eq. (IV.37) can be simplified to

$$\phi(j \omega) \approx \arctan \left( \frac{\omega}{\omega_{oi} Q_{oi}} \right)$$  \hspace{1cm} (IV.38)

This results in the delay of a biquad being

$$\tau = \frac{d\phi}{d\omega} = \frac{\omega_{oi} Q_{oi}}{\omega_{oi}^2 Q_{oi}^2 + \omega^2}$$  \hspace{1cm} (IV.39)

If a cutoff frequency is so large that $(\omega_{oi} + Q_{oi})^2 \gg \omega^2$ is valid, Eq. (IV.39) can be further simplified to

$$\tau = \frac{1}{\omega_{oi} Q_{oi}} = \text{constant}$$  \hspace{1cm} (IV.40)

In each biquad, the condition $(\omega_{oi} + Q_{oi})^2 \gg \omega^2$ may not be true, but Eq. (IV.40) still gives a hint that a larger cutoff frequency will result in a better linearity of the phase in a fixed frequency range. In the other words, a larger cutoff frequency will lead to a reduction in the change of the delay.

Based on the starting circuit capacitors in Eq. (IV.35) and keeping the above considerations in mind, the optimum values are obtained through SPICE simulations. The
final values are $C_{11,r} = 4 \, \text{pF}$, $C_{12,r} = 4 \, \text{pF}$, $C_{13,r} = 0.8 \, \text{pF}$, $C_{21,r} = 0.8 \, \text{pF}$, $C_{22,r} = 0.8 \, \text{pF}$, and $C_{23,r} = 0.8 \, \text{pF}$. These values yield $Q_{o1} = Q_{o2} = 0.588$ and $Q_{o3} = 1.171$, which are close to the values shown in Eq. (IV.31b). Also, these values result in the cutoff frequencies for every biquad being larger than the values shown in Eq. (IV.31a).

In the process being used here, only a unit capacitor is available ($0.25 \, \text{pF}$ each). This requires a further adjustment of $0.75 \, \text{pF}$ for $C_{13,r}$, $C_{21,r}$, $C_{22,r}$, and $C_{23,r}$ and $4.0 \, \text{pF}$ for $C_{11,r}$ and $C_{12,r}$.

With these capacitors, the simulation results in Fig. 51 show the gain, negative phase, and delay. For DC current, $I_s$, being $120 \, \mu\text{A}$, the -3 dB frequency is at least 4 MHz, the linear phase reaches to more than 5 MHz, and the delay is near 86 nS which is kept until 3 MHz. When $I_s$ changes from 90 $\mu\text{A}$ to 150 $\mu\text{A}$ for tuning, the delay is adjusted from 79 nS to 93 nS, but some small delay drop occurs at lower bias currents.
Figure 51. SPICE simulations of the gain <1>, delay <2> and phase <3> responses of the filter in Fig. 49 for Is values of 90 \( \mu \)A (solid line), 120 \( \mu \)A (dashed line), and 150 \( \mu \)A (dotted line).
CHAPTER V

ASP CIRCUITS WITH VERY-HIGH-FREQUENCY RESPONSE

In the last chapter, two high-frequency filters have been designed as applications of the OTAs and as tools to test the performances of the OTAs. Also, the successes of these designs show that filters operating in megahertz frequency range can be easily designed in transconductance-C structures. Note that, currently, it is difficult to design these filters with other techniques such as the switched-capacitor method. The successes of these designs also show that with a good transconductance element or OTA an analog signal processing circuit can be easily designed.

The next question is what is the frequency limitation of a transconductance-C system? Is it the bandwidth of the transconductor or the technique itself? Our primary prediction is that a transconductance-C system can use the whole frequency range provided by the transconductance element. In order to prove this assumption, a transconductance element with higher bandwidth should be available. So, in this chapter, we will try to design an OTA with gigahertz bandwidth and to build some simple transconductance-C systems based on this OTA.

There are two ways to design an OTA with very-high bandwidth. One is to use AC compensation/negative feedback techniques, and another is to use faster available devices. Generally, the former method is widely used in the design of op-amps (operational amplifiers) which have very high open-loop gain and poor bandwidth. Because the product of gain and bandwidth is a constant, negative feedback can improve the bandwidth by reducing the gain. But, in an OTA, there is only one amplifying stage, so the open-loop gain is relatively small. Also, and most important, in an OTA the interest-
ing output parameter is current, not voltage, so the output parameter is, ideally, independent of the output load. If a small output resistance is used to simulate the short circuit for obtaining an accurate output current, or transconductance of the OTA, the equivalent gain of the OTA is very small. Therefore, the bandwidth is very large. This is the reason why a current-mode amplifier has much better bandwidth than a voltage-mode amplifier. Negative feedback usually does not improve the bandwidth in OTAs. Also, this kind of improvement, eventually, is limited by the the relatively low intrinsic unity-gain frequency, \( f_T \), of CMOS transistors. So, in this thesis, in order to improve the bandwidth by an order of magnitude, the AC compensation technique will not be used. Transistors with a larger \( f_T \) than CMOS transistors will be used to extend the bandwidth of OTAs. AC compensation techniques will be used only for the compensation of excess phase shift (see Sections II.3.3.A and V.1.1).

As mentioned in the introduction, both GaAs MESFETs and silicon bipolar transistors have an \( f_T \) larger than gigahertz. Because GaAs technology was new and immature when we started this research, we prefer to design the first GaAs OTA to see whether this technology is suitable for transconductance-C circuits or not.

In Section V.1, a GaAs OTA with 7 GHz cutoff frequency will be designed as a building block. With this OTA, a high-frequency inductor can be designed with optimization. Furthermore, with the simulated resistor, inductor, and capacitance, a high-order filter can be built.

V.1. A 7 GHz GaAs OTA

At the time the design of the GaAs OTA was started, no GaAs OTA was designed or published yet. Therefore, there was no guide to follow. The design issues are: (i) There are only MESFETs instead of MOSFETs available, so the input range is limited by a Schottky diode, (ii) There is usually only an N-channel device, so many design tech-
niques cannot be used because they need the complementary devices, and (iii) the output impedance of GaAs MESFET is so low that it cannot meet the basic requirement for an OTA.

In Section V.1.1, the linearity of a differential pair with source degeneration will be discussed. Also, the compensation of excess phase shift will be mentioned. From this simple configuration, a new method for the enhancement of output impedance will be presented in Section V.1.2. Then, a new tuning method will be designed in Section V.1.3 to overcome the limitations of Schottky diodes and the availability of only N-channel devices. In the remaining sections, the other design considerations, such as DC level shifting, common-mode feedback, temperature variations, and bonding wire inductances, will be discussed. Simulation results will be shown.

V.1.1. Linearity

For 1 \( \mu \)m GaAs depletion-mode MESFETs developed by TriQuint Semiconductor, Inc., the drain current, \( I_d \), can be expressed as [219]

\[
I_d = k_1 \left( \frac{W}{L} \right) (k_2 V_{gs} - V_T)^2 (1 + \lambda V_{ds})
\]  \hspace{1cm} (V.1)

where \( k_1 \) is a transconductance parameter, \( k_2 \) is a constant, and all other parameters have the same definitions as those of MOSFETs which were defined before. From Eq. (V.1), it can be seen that the V-I characteristic of a MESFET is very close to that of a MOSFET, except \( k_2 \neq 1 \). Note that, late, TriQuint Semiconductor, Inc. had a new model for GaAs MESFET [229], but it was not used in this thesis because the fabrication of this GaAs OTA was started already.

If this kind of MESFET is used in a basic differential pair shown in Fig. 52 with \( R_E = 0 \), the transfer function, assuming \( \lambda = 0 \), can be expressed via:

\[
V_{gs1} - V_{gs2} = \frac{1}{k_2} \left[ \frac{I_{d1} L_1}{k_1 W_1} \right]^{1/2} - \left( \frac{I_{d2} L_2}{k_1 W_2} \right)^{1/2}
\]  \hspace{1cm} (V.2)
For a matched differential pair, \( W_1 = W_2 = W \) and \( L_1 = L_2 = L \), with the definition that \( k = k_1 k_2^2 \) (\( W/L \)), \( V_i = V_{gs1} - V_{gs2} = V_1 - V_2 \), and \( I_o = I_{d1} - I_{d2} \), the output current is

\[
I_o = \begin{cases} 
2I_D & V_i \geq \left( \frac{2I_D}{k} \right)^{1/2} \\
2\sqrt{\bar{k}I_D} V_i \left( 1 - \frac{kV_i^2}{4I_D} \right)^{1/2} & \left| V_i \right| \leq \left( \frac{2I_D}{k} \right)^{1/2} \\
-2I_D & V_i \leq -\left( \frac{2I_D}{k} \right)^{1/2} 
\end{cases} \tag{V.3a}
\]

and the transconductance

\[
g_m = \sqrt{8kI_D \left[ 2(1 - \frac{kV_i^2}{4I_D}) \right]^{1/2} - 2(1 - \frac{kV_i^2}{4I_D})^{1/2}} \tag{V.3b}
\]

Equation (V.3) shows that the expressions for drain current and transconductance are the same for both MESFET and MOSFET except for the different values of \( k \). Then, the input range for the transconductance with one percent nonlinearity is expressed as:

\[
\left| \Delta V_i \right| \leq 0.14 \sqrt{4I_D/k} \tag{V.4}
\]
In GaAs MESFETs, the "knee" between the saturation and nonsaturation regions is "softer" than that in MOSFETs [219], so \( k_2 = 0.5 \) is smaller than that for MOSFETs which has \( k_2 = 1 \). This results in the linear range defined in Eq. (V.4) for MESFETs being four times larger than that for MOSFETs. For example, a typical \( |\Delta V_i| \) for GaAs MESFETs is 0.15 V, compared to 0.07 V for CMOSFET shown in Section II.1.3.

Due to the above reason, a simple and effective technique, source degeneration, is chosen to increase the linear input range of the GaAs OTA. Another and important reason for using this simple technique is that the input range of a MESFET is limited by the Schottky diode, which will be discussed in detail below. So, to use the techniques discussed in Chapter II in GaAs technology does not help.

By considering the effect of \( R_E \) as local negative feedback \((R_E \neq 0)\), the global transconductance, \( G_m \), of total circuit can be derived as:

\[
G_m = \frac{g_m}{1 + g_m R_E}
\]  
(V.5)

with \( g_m \) defined in Eq. (V.3b).

Equation (V.5) shows that a small change \( \Delta g_m \) results in a corresponding change \( \Delta G_m \); the relationship between \( \Delta G_m \) and \( \Delta g_m \) is

\[
\Delta G_m = \frac{\Delta g_m}{(1 + g_m R_E)^2}
\]  
(V.6)

and shows that the change \( \Delta G_m \) is always less than the change \( \Delta g_m \) because the term \((1 + g_m R_E)^2\) is always greater than 1. This means that the linearity is improved by a factor \((1 + g_m R_E)^2\). The larger the value of the resistor \( R_E \), the more linearity is improved. This conclusion is verified by SPICE simulations performed on a typical 1 \( \mu m \) depletion-mode GaAs MESFET with \( V_T = -0.35 \) V and \( f_T = 12 \) GHz. Note that here \( V_T \) is an internal threshold voltage and the external threshold voltage \( V_T \) is equal to \( V_T / k_2 = -0.7 \) V. The results in Fig. 53 show that, when \( R_E \) increases from 100 \( \Omega \) to 500 \( \Omega \), the linear dif-
Figure 53. SPICE simulation results for the output current as a function of the resistor $R_E$ in the circuit in Fig. 52.

**Differential** input range increases from 0.4 V to 0.8 V.

It is well known that between the gate and source terminals of a MESFET there exists a metal-semiconductor contact which forms a Schottky diode with a forward conducting voltage $V_{TD}$. To guarantee that this diode is off, the gate to source voltage $V_{gs}$ of a MESFET should be restricted to:

$$V_{gs} \leq V_{TD}$$  \hspace{1cm} (V.7)

For a depletion-mode MESFET with $V_T = -0.7V$ and $V_{TD} = 0.7V$ as a typical value for a GaAs Schottky diode, the range for $V_{gs}$ is approximately 1.4 V. Then, 0.8 V linear input range of the GaAs OTA shown in Fig. 53 reaches more than half of this limit. Considering to leave some voltage margins for a change in DC operating point for the purpose of tuning, the range of 0.8 V is relatively large. Also, for the special applications with very-high-frequency response, 0.8 V is considered satisfactory.
A further reason for choosing the source degeneration technique in this design is that, in conjunction with one capacitor, the resistors $R_E$ provide one zero to be used for AC compensation. Due to the simple models used in the small-signal analysis for calculating $G_m$, Eq. (V.5) is only a simplified equation. In practice, there exist nondominant high-frequency parasitic poles, which result in small, but nonnegligible excess phase shifts. The effect of the total excess phase shift is commonly modeled by multiplying Eq. (V.5) by a term $e^{-st}$. When the condition $|s\tau| \ll 1$ is true, $e^{-st}$ can be simplified as follows:

$$e^{-st} \approx \frac{1}{1 + s\tau} \quad \text{(V.8)}$$

By using (V.8), Eq. (V.5) for $G_m$ can be written as:

$$G_m \approx \frac{g_m}{(1 + g_m R_E)(1 + s\tau)} \quad \text{(V.9)}$$

Equation (V.9) shows that an excess phase shift affects the AC performance of the transconductance $G_m$.

In order to reduce the excess phase shift, a capacitor $C_E$ is connected in parallel with resistor $R_E$. Then, the global transconductance becomes

$$G_m \approx \frac{g_m (1 + sC_E R_E)}{[1 + (g_m + sC_E R_E)(1 + s\tau)]} \quad \text{(V.10)}$$

From Eq. (V.10), it can be observed that one zero is introduced by $C_E$. If $C_E = \tau/R_E$ is chosen, the term $(1 + s\tau)$ is canceled by the term $(1 + sC_E R_E)$ as expected. Due to the small value of $\tau$, $C_E$ will be so small that the effect of the term $|sC_E|$ in the denominator of Eq. (V.10) can be neglected compared to $g_m$ up into the 10 GHz frequency range. Also, a small capacitance $C_E$ is easily fabricated in IC processing.

The performance of the compensation capacitor $C_E$ will be verified by SPICE simulations given in Section V.1.6.
V.1.2. Output Impedance

The main advantage of GaAs MESFETs is high speed which is contributed by both the GaAs material and the short channel length $L$ of the devices. But small $L$ introduces a relatively large channel-length-modulation parameter $\lambda$, typically $\lambda = 0.3V^{-1}$ for $L = 1\mu m$, and, in turn, results in a higher output conductance, $g_o = \lambda I_d/(1 + \lambda I_d)$. This disadvantage makes it impossible to directly apply GaAs technology to a transconductance circuit and calls for special circuit design techniques to obtain a high output impedance.

One of the most popular techniques used for MOSFET or Bipolar technologies is the "cascode stage" [260] which is very effective for gain enhancement by increasing the total output impedance of the circuit. But, unfortunately, this technique does not apply to GaAs technology, because it needs complementary devices such as P- and N-channel transistors to form both drive and load in cascode form and, to date, only N-type devices are available in GaAs technology. A different technique, "self-bootstrapped" [238], is suitable for circuits with single-channel devices; it has a simple circuit structure similar to "cascode." The only problem with applying this technique to GaAs technology is that the frequency response is reduced.

A typical amplifier with cascode driver and "self-bootstrapped" load is shown in Fig. 54a; it is restricted by

$$I_{d1} = I_{d2} \quad (V.11a)$$

and

$$V_{ds1} = -V_{gs2} \quad (V.11b)$$

From Eq. (V.11), two conclusions can be drawn. First, Eq. (V.11) provides negative feedback: When $I_{d1}$ increases, $V_{ds1}$ increases because $V_{gs1} = 0$. From Eq. (V.11b), $V_{gs2}$ decreases (becomes more negative); this results in $I_{d2}$ or, from Eq. (V.11a), $I_{d1}$ decreasing. With this negative feedback, the current $I_{d1}$ trends to be a constant and, therefore,
the output impedance is large. Second, the width of Q2 must be much larger than that of Q1 to satisfy Eq. (V.11a), because $V_{gs1}$ is zero and $V_{gs2}$ has a negative value.

The equation for the total output impedance, $Z_A$, of the "self-bootstrapped" load is

$$Z_A \approx \frac{R_A}{1 + s R_A C_A}$$

(V.12a)

where $R_A$ and $C_A$ are total output resistance and capacitance, respectively, which can be expressed [212] as functions of transconductance $g_m$, output conductance $g_{ds}$, and the parasitic capacitors, $C_{gd}$, $C_{gs}$, and $C_{ds}$; with $\{C_{gs}, C_{gd}\} \gg C_{ds}$:

$$R_A \approx \frac{g_m^2}{g_{ds1} g_{ds2}} \text{ and } C_A \approx C_{gd2}$$

(V.12b)

From Eq. (V.12), it follows that the frequency response of $Z_A$ depends on $C_{gd2}$, which, therefore, must be as small as possible. Unfortunately, the second conclusion reached above points out that the width of Q2, $W_2$, is usually larger than that of Q1, $W_1$; therefore, a large capacitor, $C_{gd2}$, is introduced. (In practice, for a depletion-mode MESFET with $V_T = -0.35V$, $W_2$ is four times larger than $W_1$). Consequently, the total output impedance decreases with increasing frequency so fast that a low impedance appears in the required high-frequency range. Thus, an improvement of the "self-bootstrapped" circuit should be considered.

A direct method for decreasing the capacitance of $C_{gd2}$ is to reduce the width of Q2 and, consequently, the widths of Q1 and Q4 in Fig. 54a. With the reduced width of Q4, $I_4$ decreases which results in a reduced transconductance. One method to maintain the transconductance is to introduce a bypass current from Vdd to the drain of Q4 by using a current source so that $I_4$ is unchanged and the current in Q2 can be reduced. In [41], a circuit called "improved cascode" gain stage was designed with this principle in NMOS technology. But, the single transistor load used there is not suitable for GaAs technology. Even if a small size of load transistor is used, the output impedance tended to
be low. So, a circuit which combines the techniques of "self-bootstrapping" and "improved cascode" circuits, is suggested in the following.

The circuit in Fig. 54b is used for OTA design. It can be seen that, if the width of Q5 is chosen to be equal to four times that of Q1, the width of Q2 is close to, but less than that of Q4 without any changes on the property of "self-bootstrapping." This means that high output impedance can be achieved without introducing a larger capacitor.

\[ g_m = \frac{g_m g_m (g_m + g_{ds3})}{g_m + g_{ds3} + g_{ds4} + g_{ds5}} \] (V.13a)

A small-signal analysis shows that, by assuming \( g_{ds1} = g_{ds3} \), the total transconductance \( g_m \) and total output impedance \( g_o \) can be expressed as follows:
\[ g_o \approx \frac{g_{ds4} + g_{ds5}}{g_{m3} + g_{ds3} + g_{ds4} + g_{ds5}} + \frac{g_{ds2}}{g_{m2} + g_{ds1} + g_{ds2}} \]  
\text{(V.13b)}

In Eq. (V.13b), the first term is contributed by the conductance looking down from the output node, and the second term is by looking up. If the width of Q5 is less than five times that of Q1, \((g_{m3} + g_{ds3}) \gg (g_{ds4} + g_{ds5})\). Based on the fact that \(g_{ds4} \approx g_{ds5}\), an approximation for Eq. (V.13) becomes

\[ g_m \approx g_{m4} \text{ and } g_o = (2-3)g_{ds4} \left( g_{ds3}/g_{m3} \right) \]  
\text{(V.14)}

Equation (V.14) shows that the circuit in Fig. 54b has similar expressions to those of a cascode stage.

It should be pointed out that the ratio of \(W_4/W_1\) must be large, but not very large, in practice, \(W_4/W_1 \ll 10\), otherwise, the current \(I_1\) will be much less than the current \(I_5\) which violates the condition \((g_{m3} + g_{ds3}) \gg (g_{ds4} + g_{ds5})\). In other words, \((g_{ds4} + g_{ds5})\) will be comparable to or even larger than \((g_{m3} + g_{ds3})\), so that \(g_m\) will be close to half of \(g_{m4}\) and \(g_o\) will be close to \(g_{ds3}\) only.

From the above description, one can conclude that a tradeoff among transconductance, output impedance, and frequency response is possible in the design. For example, if the above ratio, \(W_4/W_1\), becomes smaller, the transconductance will increase and the frequency response will decrease. Also, it is possible to optimize the output impedance by changing the ratio.

V.1.3. Tuning Capability

As mentioned before, tuning is a basic requirement for OTA design to counter IC process tolerances and the changes of operating conditions. For the CMOS OTAs designed in Chapter II, tuning is realized by changing the DC bias current, \(I\). Generally, this method is also suitable for most of OTAs designed in CMOS or Bipolar technologies. But, unfortunately, it cannot be applied to GaAs MESFET technology, because
short and single channel devices are used. First, since devices with a short channel length are used in GaAs technology, a larger channel-length-modulation parameter $\lambda$ exists. Therefore, the effect of the voltage $V_{ds}$ on the current $I_d$ cannot be ignored as is usually done in CMOS technology. Consequently, the DC operating points can not be well-controlled when the bias current $I$ changes. Second, and most important, since only N-channel devices are usually available, a "self-bootstrapping" technique should be used for output impedance enhancement as discussed in last section. The DC operating current in Q1, $I_1$, (see Fig. 54a) is approximately fixed for a given width of Q1, because the voltage $V_{gs1}$ is fixed. This means tuning cannot be performed by changing the gate voltage of Q1 or Q2 as is usually done in a "cascode" structure. If the gate of Q1 is not connected to its source, the condition in Eq. (V.11b) is invalid and the negative feedback loop is broken. This results in a decrease of the output impedance. For the circuit in Fig. 54b, it is also impossible to disconnect the gate and source of Q5 for the purpose of changing the DC current $I_5$ by varying the gate voltage of Q5. The reason is that this step would result in $V_{gs5}$ changing with the AC input voltage $V_{in}$ varying and, in turn, in a small output impedance. The conclusion here is that the currents in Q1, Q2, and Q5 in Fig. 54b cannot be changed without changing the widths of Q1, Q2, and Q5. So, a new method for tuning will be considered with the restriction that the sum of the DC currents $I_1$ and $I_5$ should be constant.

A possibility to avoid the above problem is to connect a current source $I_b$ in parallel with Q4 (see Fig. 55a). Then the relationship between $I_1$, $I_5$, and $I_4$ becomes:

$$I_4 = I_1 + I_5 - I_b$$

Equation (V.15) shows that the current $I_4$ can be changed by varying the current $I_b$ only; therefore, $I_1 + I_5$ can be kept constant. This means that the transconductance (depending on $I_4$) can be tuned without changing $I_1 + I_5$. 
In practice, a transistor Qb is used to replace the current source \( I_b \) as shown in Fig. 55b, and the current in Qb, \( I_b' \), is controlled by the gate voltage \( V_a \) of Qb. When \( V_a \) increases, \( I_b \) increases and \( I_4 \) decreases, which results in a decrease in \( g_m \). Note that, although the relationship between \( V_a \) and \( I_b' \) restricted by Eq. (V.1), is nonlinear, there is no problem for tuning. The reason is that the requirement for tuning is continuity, not linearity.

![Figure 55](image.png)

**Figure. 55.** Transconductance tuning by using shunt current source (a) or transistor Qb (b).

As mentioned in Section V.1.1, a Schottky diode exists between the gate and source terminals of a MESFET. Therefore, the range of voltage change for \( V_{gs} \) of Qb is limited by Eq. (V.7) to near 1.4 V. For the linearity consideration the range of voltage changes for \( V_{gs} \) of the active transistor Q4 will be less than 0.75 V, because the soft "knee" of a MESFET results in the transistor operating far away from \( V_T = -0.7 \) V, usually at \(-0.2 \) V. Then, if the widths of Q4 and Qb are chosen to be the same to obtain the
condition $I_4 = I_b$, the maximum range for changing $I_4$ is typically a factor of four. In practice, during tuning, the linear input range is reduced because the current of an OTA becomes saturated earlier (see the curve $I_{bias} = 80 \mu A$ in Fig. 24). In order to avoid this situation in the GaAs OTA, the maximum tuning range is further reduced. The reason to avoid this from happening in GaAs OTA, not in CMOS OTAs, is that the input range of the GaAs OTA is relatively small. If the width of Q4 is greater than that of Qb, the maximum tuning range is smaller. In the opposite way, the range may be larger. But, due to the restrictions by Eq. (V.7) and by the linearity of the OTA, the maximum tuning range is typically restricted to a factor of four for this configuration. Otherwise either $V_{gs4}$ is out of the range defined by Eq. (V.7) or the linearity is poor. So, the conclusion is that the possible tuning range with acceptable linearity is much less than a factor of four, say a factor of three.

In order to verify the above conclusion, SPICE simulations were performed on the circuit in Fig. 52 with the transistors Q1 and Q3 (Q2 and Q4) being replaced by the circuit in Fig. 55b. The results are in Fig. 56 in the curves labeled $V_c = -5 V$ ($V_c$ is identified below in Fig. 57). The curves show that, for $W_4 = W_b$ and an acceptable linearity, tuning caused by changing $V_d$ gives a relatively small range of a factor of 1.8. There are two reasons for this small tuning factor: First, the current $I_{out}$ in Fig. 56 is the total differential current, not the DC current in Q4. Then, the change of $I_{out}$ not only depends on the current in Q4, but also on the resistor $R_E$. Eq. (V.5) shows that the changing of $I_{out}$ which determines $G_m$ is less than that of $I_4$ which determines $g_m$. The purpose of $I_{out}$ used is that the interesting parameter is $I_{out}$, not $I_4$, because the total transconductance is determined by $I_{out}$. Second, although $V_d$ in Fig. 56 is changed from -0.8V to 0.3V, $V_{gsb}$ changes, actually, only from around -0.4 V to +1.5 V. The reason is that, when the circuit in Fig. 55b is used as an input stage in a differential pair, the source terminals of the transistors Q4 and Qb are connected to a current source, $2I_D$ in Fig. 52. So, the node of
the common source terminal is an AC ground or a floating node. When the DC voltage \( V_a \) changes, both \( V_{gs4} \) and \( V_{gsb} \) change. This results in the change of \( V_a \) being shared by Q4 and Qb. An accurate voltage for the common source terminal can be calculated as follows.

\[
\begin{align*}
V_c &= \frac{W}{L} \left( k_2 (0 - V_s) + V_T \right) + k_1 \frac{W}{L} \left( k_2 (0.3 - V_s) + V_T \right)^2 = 2k_1 \frac{W}{L} (V_T)^2 \\
\text{(V.16a)}
\end{align*}
\]

In Eq. (V.16a), \( V_s \) is the voltage of the common source terminal. Consider \( k_2 = 0.5 \) and \( V_T = -0.35 \) V, then \( V_s = 0.166 \) V.

In another case, \( V_a = -0.8 \) V which is so low that Qb can be considered to be off. So, all current flows through Q4, and the new equation is
Transconductance tuning by using the combination of "digital" and "analog" techniques.

\[ k_1 \frac{W}{L}(k_2(0-V_s)+V_T)^2 = 2k_1 \frac{W}{L}(V_T)^2 \quad (V.16b) \]

From Eq. (V.16b), \( V_s \) can be calculated as -0.3 V. So, the actual range of the \( V_{gsb} \) change is from -0.5 V to +0.134 V and the range of the \( V_{gs4} \) change is from -0.166 V to 0.3 V; both of these ranges are within the one predicted by Eq. (V.7).

To enlarge the tuning range, two additional transistors Q6 and Q7 are added to the circuit in Fig. 55b as shown in Fig. 57. Here, Q7 is a pass gate transistor controlled by the voltage \( V_c \), and Q6 is another possible active path for the input signal. When \( V_c = -5V \), both Q7 and Q6 are off. Then, the input signal is applied to Q4 only. Thus, the "equivalent" width of the input device for the AC signal is \( W_4 \) which determines the transconductance \( g_m \). If the widths of Q4 and Qb are chosen to be the same, the tuning range is less than a factor of three as discussed above. When \( V_c = 0 \), Q7 is on and, in turn, Q6 is on. The input signal now is applied to both Q4 and Q6, which results in the
"equivalent" width of the input devices equal to \( W_4 + W_6 \). Then, for \( W_4 = W_6 = W_b \), the "equivalent" active current increases from \( 1/2 \) to \( 2/3 \) of the total current or increases by 66% and can be changed by varying \( V_a \) in a new range different from that for \( V_c = -5 \ V \). This results in the tuning range being doubled if the two different ranges do not overlap. But in practice the ranges should overlap a little, because tuning should be continuous. If \( W_4 = W_6 = W_b \), this condition is always true because the AC current only can increase by 66% which is less than the tuning factor of 80%. For example, if the total current is chosen as unity, the "active" current is 0.5 for \( V_c = -5 \ V \). With a tuning range as a factor 1.8, the current change is from 0.357 to 0.643. When \( V_c = 0 \ V \), the current increases from 0.5 to 0.83 and can be changed from 0.59 to 1.06. Then, the total tuning range is continuously changed from 0.357 to 1.06, increasing from the original factor of 1.8 to a factor of 2.97.

The function of tuning has been checked by SPICE simulation with \( W_4 = W_b = 7 \ \mu m, W_6 = 6 \ \mu m \), and the load resistance equal to 1 k\( \Omega \). Because these three widths are almost the same, a tuning range is continuous. The results in Fig. 56 show that when only \( V_a \) changes the transconductance can be tuned continuously by nearly a factor of 1.8 in a good liner range for \( V_c = 0 \ V \) or \( V_c = -5 \ V \). Combining these two ranges, the tuning range is nearly, but less than, doubled and is still continuous. This conclusion also can be seen from Fig. 60a which shows the tuning range and, also, indicates an important fact: the bandwidth of the OTA is nearly independent of tuning.

Because Q7 is used as a switch and \( V_c \) has only two possible values: 0 and -5V, active current change caused by changing \( V_c \) may be called "digital tuning." Correspondingly, the change caused by changing \( V_a \) is called "analog tuning," because the change is continuous. It is obvious that, if a larger tuning range is required, further transistor pairs, such as Q6 and Q7 can be used. This means that the scheme can be extended by using additional transistors Q6i, addressed by corresponding pass-transistors Q7i and con-
trolled by voltages $V_{c1}$ through programming. In this manner, a combination of digital and analog tuning is employed and a large tuning range can be obtained.

### V.1.4. DC Level Shifter

As mentioned before, in GaAs MESFET technology, only N-channel devices are usually available. Thus, if the DC input voltage level of the OTA is zero, the DC output voltage level is greater than zero. Therefore, a level shifter must be used because direct coupling is usually used for integrated circuits.

As a conventional level shifter, a source follower Q2 in Fig. 58 shifts the voltage by one $V_{gs2}$. An n-stage multiple source follower, therefore, can shift the voltage by

$$
\sum_{i=0}^{i=n} V_{gs,i}
$$

Unfortunately, source followers can not be used in a transconductance circuit, because the output impedance of a source follower is very low, which, in turn, results in a low output impedance for the designed OTA. Also, in a transconductance circuit, the interesting output parameter is current, not voltage. So, if a source follower is employed in an OTA, the output current, for example $I_{d2}$ in Fig. 58, becomes a complicated function to input voltage $V_{in}$:

$$
I_{d2} = k' [k_1 (I_{d1} r_{ds1} + I_{d2} r_{ds2} + 2V_{ss}) - V_T]^2
$$

(V.17)

where $I_{d1}$ is a function of $V_{in}$ as shown in Eq. (V.1), $k' = k_1 \frac{W_2}{L_2}$, and $r_{ds1}$ and $r_{ds2}$ are the output impedances of Q1 and Q2, respectively. From Eq. (V.17), it can be observed that the output current, $I_{d2}$, is an implicit function of $V_{in}$. The transconductance of the total circuit, therefore, is hard to be predicted and is not a constant as required.

In order to maintain the properties of the transconductance circuit, a level shifter consisting of only the diodes $d_1$ to $d_{14}$ in Fig. 59 [97] is used. The output current is approximately unchanged after passing the diodes. Because the diodes are directly connected to the output stage, or to the transistors Q7 and Q8 in Fig. 59, a parasitic
impedance, $nR_d/(1 + sC_dR_d)$, is connected in series between the output stage and the load. Where, $n$ is the number of diodes used, $C_d$ is the diode capacitance, and $R_d$ is the diode parasitic resistance. The diode parasitic resistance degrades the frequency responses of the OTA in the high-frequency range, so two larger bypass capacitors $C_3$ and $C_4$ with the value of 0.9 pF are used. In Section V.1.6, the SPICE simulation results will show that this design gives an acceptable frequency response with the required performance of the DC level shifter.

V.1.5. Common-Mode Feedback Circuit

In order to achieve the advantages of a fully balanced circuit structure, a differential output stage is used in this GaAs MESFET OTA. As discussed in Chapter III, a common-mode feedback circuit must be used to stabilize the DC output level. It is also discussed in Chapter III that a CMF circuit is possible to cause DC offset during tuning. Fortunately, tuning of this OTA is not implemented by changing the bias current $I_g$, but by varying the current in $Q_b$ in Fig. 57. Then the DC current in the output stage does not change with tuning, which results in constant DC voltages at the output nodes. Thus, the popular and simple CMF circuit shown in Fig. 28(1) can be used. The modification made here is to add the level shifting diodes $d_{15}$ to $d_{23}$ as shown in Fig. 59. The purpose of
Figure 59. Circuit diagram for a GaAs OTA.
using \( d_{20} \) to \( d_{23} \) is to generate an appropriate voltage drop to guarantee that Q26 works in the saturation region.

The simulation results for the common-mode feedback circuit will be shown in the next section.

V.1.6. Simulation Results

The OTA with the stages discussed above is drawn in Fig. 59. From the circuit diagram, it can be observed that the input stage is a cascode circuit. The transistors Q1 to Q6 and Q15 to Q16 are used for input and tuning, and the transistors Q7 to Q14 form an improved circuit for output impedance enhancement. The source degeneration is formed by the resistors \( R_1 \) and \( R_2 \) with the value of 500 \( \Omega \) each, and the capacitors \( C_1 \) and \( C_2 \) with the value of 50 \( \mu F \) each are used for reducing the excess phase shift. The diodes d1 to d14 are the level shifter and the capacitors \( C_3 \) and \( C_4 \) with the value of 0.9 \( pF \) are used for AC bypass.

Note that the sum of the widths for Q1, Q3, and Q5 is the same as the sum of widths for Q7 and Q12, but, the width of Q7 is not the same as that of Q9. The reason is that the current in Q9 is the sum of the currents of Q7 and Q21, rather than that of Q7 only. Another reason is that \( V_{gs9} \) is equal to zero while both \( V_{gs7} \) and \( V_{gs21} \) may not be zero.

The OTA has been simulated by SPICE with the parameters used in Section V.1.1 and the load resistance of 1 k\( \Omega \). The widths for all transistor are shown in the circuit with the unit "\( \mu m \)", and the power supplies are +5 V, +10 V, and -5 V. \( V_a \) changes from -0.8V to 0.3V as discussed in Section V.1.3.

The simulation results for frequency response in Fig. 60a show that this OTA has a very-high bandwidth: \( f_{-3db} \) is at least 7 GHz over the whole tuning range. The results for linearity are shown in Fig. 56, which has been discussed already. The output
impedance shown in Fig. 60b is around 160\(k\Omega\) at low frequencies and is larger than 100 \(k\Omega\) until \(f = 100\) MHz. If required, the output impedance can be further increased by decreasing the transconductance as a tradeoff. The output parasitic capacitor is as low as \(-16\) \(fF\), and it can be reduced by decreasing the sizes of the transistors. The input capacitor is also 16 \(fF\) with a very high input resistance.

The common-mode feedback circuit has also been tested; the simulation results in Fig. 60b show the common-mode output impedance of the OTA is only 5 \(k\Omega\). The comparison of this value to the low-frequency differential output impedance of 160 \(k\Omega\) shows that a high common-mode rejection ratio is achieved.

In order to test the performance of compensation capacitors \(C_1\) and \(C_2\) in Fig. 59, an integrator is built by this OTA terminated by a grounded capacitor, \(C_L = 1\) \(pF\) (see Fig. 3b). The simulation results in Fig. 61 show that for \(C_1 \leq 150\) \(fF\) the excess phase shift can be totally compensated up to \(f = 1\) GHz without leading to any peaking in the total output current (see Fig. 60a).

In practice, it is impossible to change the value of \(C_1\) directly for phase tuning as discussed in Chapter II. One possible choice for making the compensation electrically variable is to place a GaAs MESFET in parallel with the capacitor \(C_1\) by connecting source and drain of the MESFET to the terminals of \(C_1\) and the gate of the MESFET to a negative voltage \(V_C\). Thereby, \(C_1\) is augmented by the voltage-dependent junction capacitor of two reverse-biased Schottky diodes \(D_{gs}\) and \(D_{gd}\) in series, so that the phase can be varied by controlling \(V_C\). Usually, the junction capacitor is a nonlinear function of the applied signal and control voltages; whether this causes difficulties in the design of a filter and the phase control loop has to be carefully evaluated.

**V.1.7. Other Considerations**

The OTA in Fig. 59 has been also simulated for different temperatures. The
Figure 60. SPICE simulation results for frequency response of the GaAs OTA with different tuning voltages (a) and output impedances (b) for <1> the differential output and <2> the common-mode output.
results in Fig. 62 indicate that the -3dB frequency changes by 12% and the low-frequency transconductance changes by only 0.5% as temperature varies from 27°C to 100°C (curves <1> and <2> in Fig. 62). These results can be explained by two facts: first, in the indicated temperature range, for intrinsic reasons, the MESFET used has approximately a 14.5% change of $f_T$ and a nearly 0.6% change of $g_m$ [219]. These variations are very close to the simulation results. Second, in the OTA design, the source degeneration used provides local feedback to stabilize the input stage to temperature variation.

Because the GaAs OTA works in the gigahertz frequency range, the parasitic inductances and resistances of bonding wires from pads to pins may become nonnegligible. So, the OTA has been simulated with 2 nH bonding inductors for all power supplies and 2 nH bonding inductors and 50 Ω wire resistors for both input and output terminals.
Figure 62. The simulation results for the output current of the OTA in Fig. 59: <1> $T = 27^\circ C$ without wire inductors, <2> $T = 100^\circ C$ without wire inductors, and <3> $T = 27^\circ C$ with 2 nH wire inductors.

The simulation result in Fig. 62 (curve <3>) shows that there is no peaking or phase lead present; there is only a 10% change in frequency response. If a smaller value of 1 nH is used for the wire inductors in simulation, no frequency variation is observed. The reason can be found in the low operating gain of the OTA with 1 k\Omega load and the very small parasitic capacitances (16 fF for both input and output capacitances).

V.2. A 1 GHz GaAs GYRATOR-C INDUCTOR

With the GaAs OTA, we can build a simulated inductor. The reason we are interested in building an inductor first is that inductors have played an important role in analog circuits for signal processing because, in conjunction with capacitors, inductors can provide complex poles. In recent years, in order to include inductors in integrated circuits, the bulky discrete inductors are replaced by simulated inductors. Particularly, with increasing requirements on the accuracy and the high-frequency behavior, the parasitic
parameters in discrete inductors are so large that simulated inductors must be used. Although using an RC-active network can avoid inductors, this network usually leads to circuits which are very much more sensitive to component tolerances than an LC network [1]. Therefore, building a simulated inductor is very desirable.

The circuit diagram of a gyrator-C inductor implemented by OTAs and a capacitor is shown in Fig. 42d. The difficulty for building a good gyrator-C inductor is the parasitics associated with OTAs. It is well known that, ideally, a gyrator-C inductor is built by two voltage-controlled current sources and one capacitor. But, when OTAs are used to replace the current sources, the input and output impedances of the OTA (see Fig. 63) form an RLC network (see Fig. 64a) which is equivalent to the inductor in Fig. 42d. These parasitics will affect the resonant frequency $\omega_0$, quality factor $Q$, and the valid range of constant inductance. One way to reduce the effect of the parasitics is to design OTAs with small input and output admittances. Another way is to find the preferred relationship among the parasitics of the OTAs, by considering which parameters are important for the design and which ones not. An optimum result can be obtained by carefully appropriate tradeoffs.

![Figure 63. Small-signal equivalent circuit for single-ended OTA with $g_m$ being the transconductance, $Y_i = g_i + sC_i$ and $Y_o = g_o + sC_o$, where $g$ is the parasitic conductance and $C$ is the parasitic capacitance.](image)

V.2.1. Design Considerations for Gyrator-C Inductors

The equivalent circuit of a gyrator-C inductor (Fig. 42d) is shown in Fig. 64a with $L$ being the inductor, $R_{eq}$ the equivalent loss-resistor, $R_p$ the equivalent parallel resistor,
and $C_{eq}$ the equivalent parasitic capacitor. Using the small-signal equivalent circuit for an OTA shown in Fig. 63, the small-signal equivalent circuit for a single-ended structure is obtained in Fig. 64b and a small-signal analysis shows that input conductance of this inductor is

$$Y_{in} = (g_{i1} + g_{o2}) + s(C_{i1} + C_{o2}) + \frac{g_{m1}g_{m2}}{(g_{i2} + g_{o1}) + s(C + C_{o1} + C_{i2})}$$  \hspace{1cm} (V.18)

where $g_{m1}$ is the transconductance of the first OTA and $g_{m2}$ the transconductance of the second OTA, and all other parameters are defined in the figure heading of Fig. 63. Note that, in Fig. 42d, the transconductances for both OTAs are the same, but different ones are assumed for Eq. (V.18).

Figure 64. A equivalent circuit for the gyrator-C inductor shown in Fig. 42d. (a) RLC equivalent network. (b) Small-signal circuit for single-ended structure. $Y_{in} = I_{in}/V_{in}$ is the equivalent input conductance.

Generally, the circuit capacitance $C \gg C_{o1}$ and $C_{i2}$, and $g_{oi} \gg g_{ii}$. Then, the
inductance and parasitic parameters of this inductor can be expressed as follows:

\[ L = \frac{C}{g_{m1}g_{m2}} \quad (V.19a) \]

\[ R_{eq} = \frac{g_{o1}}{g_{m1}g_{m2}} \quad (V.19b) \]

\[ C_{eq} = C_{i1} + C_{o2} \quad (V.19c) \]

\[ R_{p} = \frac{1}{g_{o2}} \quad (V.19d) \]

Further, the self-resonance frequency, \( \omega_{o} \), self-resonance quality factor, \( Q_{o} \), and inductor quality factor, \( Q_{L} \), can be calculated as

\[ \omega_{o} = \left( \frac{1}{LC_{eq}} \right)^{1/2} = \left( \frac{g_{m1}g_{m2}}{C(C_{o2} + C_{i1})} \right)^{1/2} \quad (V.20a) \]

\[ Q_{L} = \frac{\omega L}{R_{eq}} = \frac{\omega C}{g_{o1}} \quad (V.20b) \]

\[ Q_{o} = R_{p}\left( \frac{C_{eq}}{L} \right)^{1/2} = \frac{1}{g_{o2}} \left( \frac{g_{m1}g_{m2}(C_{o2}+C_{i1})}{C} \right)^{1/2} \quad (V.20c) \]

From Eq. (V.20), it can be observed that the different parasitic parameters of OTAs do play different roles in determining the parameters of the inductor. For example, the values of the capacitors \( C_{i2} \) and \( C_{o1} \) are relatively arbitrary, because they are not shown in the above equations. Therefore, there are no restrictions on \( C_{i2} \) and \( C_{o1} \). Also, the individual transconductances \( g_{m1} \) and \( g_{m2} \) are not important for the design but their product is. Similarly, only the sum of the capacitors \( C_{i1} \) and \( C_{o2} \) is relevant.

For a required inductance \( L \), in order to obtain higher resonance frequency \( f_{o} \) it can be seen from Eq. (V.20a) that the only choice is to reduce the parasitic capacitor, \( C_{eq} \). This, in turn, requires the input capacitor, \( C_{i1} \), and the output capacitor, \( C_{o2} \), to be small. Since \( C_{i1} \) can be reduced only by decreasing the sizes of transistors used in the first OTA, \( g_{m1} \) will be reduced also. But, as mentioned above, the interesting parameters
are not $g_{m1}$ or $g_{m2}$, but their product. So, a small $g_{m1}$ will not cause any problems if the product is maintained constant by increasing $g_{m2}$. Of course, increasing $g_{m2}$ will cause both capacitors $C_{i2}$ and $C_{o2}$ to increase. Fortunately, a larger $C_{i2}$, being not very large, is not important for the design and $C_{o2}$ can be reduced by employing circuit design techniques, for example, by increasing the width of transistor $q_{13}$ and decreasing the width of transistors $q_7$, $q_9$, and $q_{11}$ in Fig. 59.

Another advantage of using small-size transistors in the first OTA is that a high inductor-Q can be obtained due to the higher output impedance of the first OTA. For example, if $W_7 = 2 \, \mu m$, the low-frequency output impedance of the OTA in Fig. 59 can be as high as 1 MΩ.

When $C_{eq}$ is reduced, the self-resonance quality factor $Q_o$ is reduced also. From Eq. (V.20c), it can be observed that $Q_o$ depends on $g_{o2}$ more than on $C_{eq}$, so that the reduction on $C_{eq}$ can be compensated easily by a very small $g_{o2}$. Because large-size transistors have been used in the second OTA, $g_{o2}$ can only be reduced by using the circuit design techniques which have been used already for reducing the capacitor $C_{o2}$. Then, $g_{o2}$ cannot be further reduced to as small a value as desired. A tradeoff, therefore, exists between $f_o$ and $Q_o$.

Because a small $g_{m1}$ will be used in the design, as discussed above, the tuning range in the first OTA may be limited by transistor characteristics. Fortunately, tuning can be implemented by changing either $g_{m1}$ or $g_{m2}$, because only the product of $g_{m1}$ and $g_{m2}$ is of interest. Thus, tuning in the first OTA can be deleted, which results in the transistors $q_3$, $q_4$, $q_5$, $q_6$, $q_{15}$ and $q_{16}$ being removed from the OTA in Fig. 59. Also, as the capacitor $C_{o1}$ is less important for the design, transistors $q_{13}$ and $q_{14}$ can be removed too. Then, the circuit for the first OTA used in our gyrator-C inductor is reduced to the one shown in Fig. 65, from which it can be seen that the number of transistors in the input stage is reduced by approximately a factor of two. Note that the sizes of the transis-
Figure 65. A simplified circuit for the first OTA used in gyrator-C inductor.
tors in Fig. 65 used here will be given in next section.

Based on the above discussion, a general design procedure can be summarized as follows [142]:

a) For a required \( L \), calculate the \( g_m \) by Eq. (V.19a) by assuming \( g_{m1} = g_{m2} = g_m \).

b) Decrease \( g_{m1} \) by a factor and increase \( g_{m2} \) by the same factor.

c) Calculate all parasitic parameters for the first OTA shown in Fig. 65 and for the second OTA shown in Fig. 59.

d) Calculate the parameters, \( \omega_o, Q_L, \) and \( Q_o \) by Eq. (V.20).

e) If \( Q_o \) is very high but both \( f_o \) and \( Q_L \) are not high enough, continue to decrease \( g_{m1} \) and increase \( g_{m2} \). If \( Q_o \) is too small, increase \( g_{m1} \) and decrease \( g_{m2} \).

f) Repeat the above procedure, until satisfactory values for \( \omega_o, Q_L \) and \( Q_o \) are obtained.

V.2.2. Practical Design

According to the design procedure described in Section V.2.1, a gyrator-C inductor was designed with 1 \( \mu \)m GaAs depletion-mode MESFETs which were used in the design of GaAs OTA early. The simple circuit in Fig. 65 is used for the first OTA and the circuit in Fig. 59 is employed for the second OTA. The gate length of the transistors used in both OTAs is \( L = 1 \mu m \). The widths of the transistors used in the first OTA are \( W_1 = W_7 = 2 \mu m, \ W_9 = 6 \mu m, \ W_{11} = 20 \mu m, \ W_{17} = W_{18} = 4 \mu m, \) and \( W_{19} = W_{20} = 3 \mu m, \) and those of all diodes are 2 \( \mu m \). For the second OTA, the widths of all transistors are \( W_1 = W_5 = W_7 = 6 \mu m, \ W_3 = 5 \mu m, \ W_{13} = 11 \mu m, \ W_9 = 10 \mu m, \ W_{11} = 45 \mu m, \) and \( W_{19} = W_{20} = 2 \mu m, \) and those of all diodes are 2 \( \mu m \). \( V_c \) and \( V_a \) in Fig. 59 for the second OTA are selected as 0 V and -0.8 V, respectively, to obtain the maximum transconductance. With this choice, \( g_{m1} = 0.05 \, mS, \ R_{out1} \approx 1 \, M\Omega, \ g_{m2} = 0.25 \, mS, \ R_{out2} \approx 300 \, k\Omega, \ C_{i1} \approx 1.5 \, fF, \) and \( C_{o2} = 6.5 \, fF \).
The designed inductor was simulated by SPICE with $C = 0.15 \, pF$ and the results are shown in Fig. 66. It can be observed that the self-resonance frequency $f_o$ is 1 GHz, which is slightly smaller than the calculated value, 1.03 GHz. Note that the equations in Eq. (V.19) are obtained for a single-ended circuit. When a fully-differential structure is used, the equations should be modified by doubling the transconductances of the OTAs. Fig. 66 also shows that $Q_L$ is 80 at $f = 100$ MHz with very high $Q_o$.

![Figure 66](image)

Figure. 66. SPICE simulated frequency response of a gyrator-C inductor.

The magnitude of the inductive impedance (Curve <1> in Fig. 66) at low frequencies indicates that the equivalent parasitic resistance is around 25 $\Omega$, which is close to the calculated value of 20 $\Omega$. The difference is caused by the assumption $g_{o1} \gg g_{i2}$ used for obtaining Eq. (V.19b). Curve <1> also indicates that $L \approx 3.2 \, \mu H$ which is also close to the calculated values.
The phase of the inductor in Fig. 66 shows that if the required phase error is less than 5 degrees, the valid range of this inductor is from 10 MHz to 1 GHz. If the required phase error can be relaxed to 45 degrees, the relatively large range from 1 MHz to 1 GHz is obtained.

V.3. A SIMULATED 200 MHz GaAs LC LADDER FILTER

In the last two sections, a 7 GHz OTA and a 1 GHz inductor have been designed in 1 \( \mu \)m depletion-mode GaAs technology with \( V_T = -0.35 \) V and \( f_T = 12 \) GHz. These designs show that the frequency responses of these basic building blocks are limited by the \( f_T \) of GaAs MESFETs, which proves the prediction we made in the introduction of this chapter. In order to further prove that this prediction is valid for a complex systems, a high-order filter will be designed in this section.

Using an OTA, a simulated resistor can be easily designed as shown in Fig. 42b. With the basic elements of \( R, L, \) and \( C \), an \( LC \) ladder filter can be built. As mentioned in Chapter IV, for the purpose of testing, an elliptic filter is a good candidate because the transmission zero of the filter, which determines the frequency response and magnitude attenuation of the filter, is much more sensitive to parasitics than the cutoff frequency. So, for this purpose and to simplify the design procedure, a third-order elliptic filter is selected as an example.

The circuit diagram of a third-order \( LC \) lowpass elliptic filter was shown in Fig. 43 and its implementation by OTAs and capacitors in Fig. 44. All OTAs used are the GaAs OTA in Fig. 59 with identical transconductances, \( g_m = 0.3 \) mS, except the first OTA, which has the value of \( 2g_m \) so that the transfer function is equal to unity at DC [131]. The control voltage of the output buffer now should be connected to \( V_{SS} (-5 \) V) instead of \(-2 \) V used in CMOS, because the transistors qs2 and qs4 are depletion-mode devices. The widths for the transistors \( q_{s1} \) to \( q_{s4} \) are chosen as 4 \( \mu \)m for saving power.
and chip area.

The design considerations for a good frequency response mentioned in Section IV.2.2 are general for any kind of technology. So, the parasitic capacitances of the GaAs OTA ($C_{in} = C_{out} = 16 \, fF$) should be included in the calculations. Considering the frequency limitation caused by parasitics and using Eq. (IV.25), the circuit capacitor $C_i$ can be calculated, for $f_o = 200 \, MHz$, as $C_1 = 0.458 \, pF$, $C_2 = 0.179 \, pF$, $C_3 = 0.489 \, pF$, and $C_L = 0.272 \, pF$. These values satisfy the requirement to be large compared to parasitics.

The filter was simulated by SPICE, and the result in Fig. 67 (curve <3>) shows that the cutoff frequency is 200 MHz as designed. The transmission zero is located near 350 MHz and the notch is deep enough. These results mean that the parasitic capacitances were well predistorted when $C_i$ was calculated. The maximum value of the cutoff frequency in Fig. 67 (Curve <1>) is 300 MHz which corresponds to the maximum transconductance $(g_m)_{max} = 0.44 \, mS$ and the minimum value is approximately 130 MHz. This indicates the cutoff frequency of this filter can be tuned by a factor 2.3.

From Fig. 67, it can also be seen that the low-frequency gain is -2 dB instead of 0 dB. The reason is that the gain of the output buffer is only 0.8, instead of 1, due to the lower $g_m \cdot r_d$ product of GaAs MESFETs. If unity gain of the filter is required for the design, the self-bootstrapping technique could be used in the output buffer or $g_{m1}$ could be increased. Fig. 67 also shows that the magnitude attenuation is more than 23 dB for all tuning.

This design further verifies the prediction made early and shows that a continuous-time transconductance-C system is a good candidate for high-frequency applications.
Figure 67. The simulation results of the frequency responses with tuning for the filter in Fig. 44, but with GaAs OTAs in Fig. 59: <1> $V_c = 0$ and $V_a = -0.8$ V, <2> $V_c = -5$ V and $V_a = -0.8$ V, <3> (nominal response) $V_c = 0$ and $V_a = 0.3$ V, and <4> $V_c = -5$ V, $V_a = 0.3$ V.
CHAPTER VI

IC PROCESSES AND LAYOUT DESIGNS

In the last several sections, the design techniques for OTAs with different considerations, such as linearity, linear input range, frequency response, and common-mode feedback circuit, have been discussed. Also, the design procedures for filters with several tens of megahertz frequency range and even several hundreds of megahertz range have been presented. The designed circuits have been primarily verified by SPICE simulations. Although the simulation results can be trusted, particularly, for simulation done with Level 3 (a semi-empirical curve-fitting) model, these results do not include the process tolerances which, in practice, can be as high as 15% and may totally destroy the performances of the circuits. So, to include the process tolerance in the design is a must.

To consider the impact of process tolerances on the circuits, currently, computer simulations are not adequate. The best choice is to built the circuits and to measure the real data. Obviously, these data contain all information we are interested in.

The first step in building a chip is to design a layout which is a graphic representatives for the physical transistors, capacitors, resistors, and interconnections. After layout, a verification tool should be used to debug the errors made during the circuit or layout designs, and to check whether the parasitics are acceptable. With a clean layout, a circuit can be fabricated.

In this chapter, the processes used for fabrication will be introduced first, then the layout tools will be briefly mentioned. Next, the design considerations of layout will be discussed and, finally, the design of the OTA and the filter layouts will be presented.
VI. DESCRIPTION OF IC PROCESSES

As mentioned in the introduction, there are three different processes used in the design: (1) MOSIS from University of Southern California, (2) DABIC from Allegro Microsystems, Inc., and (3) QED from TriQuint Semiconductor, Inc. They will be introduced one by one.

VI.1. MOSIS

MOSIS (MOS Implementation Service) is a group of vendors operated by the Information Sciences Institute of the University of Southern California under the sponsorship of the U.S. Department of Defense Advanced Research Project Agency (DARPA) and the National Science Foundation (NSF). The feature of MOSIS is that it combines several projects on a single wafer fabrication run, so the cost can be reduced. In particular, the funding for fabrication for University projects can be obtained from NSF, and this is why we selected MOSIS for our fabrication.

MOSIS supports four different technologies: NMOS, CMOS/bulk, CMOS/SOS, and PCB, also MOSIS supports both digital and analog processes. In particular, MOSIS provides the scalable CMOS/bulk technology. The main advantage of this technology is that the size of channel length of devices is scalable from 3.0 µm, 2.0 µm, 1.6 µm, to 1.2 µm, because the same set of design rules covers all four feature sizes. This means that a 3.0 µm layout can be fabricated at 2.0, 1.6, or 1.2 µm by scaling a large fraction of the layout instead of redesigning. For packaging, MOSIS supports 28, 40, 64, 84, 108, or 132 pins packages.

The process used for fabrication was a 2 µm analog N-Well CMOS process with vendor’s name ORBIT and running name N11B. The process possesses two layers of metal, metal1 and metal2, for interconnections and has two layers of poly, poly1 and poly2. Generally, poly1 is used for the gates, and poly2 for interconnections. The double
poly provides a large and accurate capacitance in a small area. The package used has 40 pins.

Details of the MOSIS process file we use for SPICE simulations are shown in Appendix A.1.

The SPICE simulation results for the transistor characteristics based on the parameters in Appendix A.1 are shown in Fig. 68. It can be seen that, for a 10 μm N-channel or a 30 μm P-channel transistor, the value of the drain current is normal, the threshold voltage is normal, but the output impedance is relatively small. So, in the OTA design, a cascode stage is used for the enhancement of the output impedance.

Note that the corners (knees) between the linear and the saturation regions shown in Fig. 68a (NMOSFET) are caused by modeling, and they do not appear in practice because of the continuity of the current derivative.

VI.1.2. DABIC

DABIC is an industrial fabrication system from Allegro Microsystems, Inc., formerly Sprague Semiconductor Group. DABIC is an analog BiCMOS process with double wells and has both 2.4 μm (DABIC V) and 3.5 μm (DABIC II) technologies. When starting to fabricate our circuits, DABIC V was not available so DABIC II was used. The package used has 68 pins, but only 42 pins were used.

The fabrication in this process was free of charge in exchange for the consultant work to the design of the transconductance-C filters which the company was interested in. So, all information for this process is company-confidential, and cannot be revealed here. But, the transistor characteristics based on SPICE simulation done here are shown in Fig. 69. From there, it can be observed that the N-channel transistor is almost perfect; normal drain current, smooth transition between linear and saturation regions, and very high output impedance. The P-channel transistor is typical.
Figure 68. SPICE simulation results for the drain current vs. drain-to-source voltage characteristics of an N-MOSFET (a) and a P-MOSFET (b) from MOSIS.
Figure 69. The simulation results for the drain current vs. drain-to-source voltage characteristics of an N-MOSFET (a) and a P-MOSFET (b) from Allegro Microsystems, Inc.
VI.1.3. QED

QED is also an industrial fabrication system provided by TriQuint Semiconductor, Inc. QED is a 1 μm GaAs MESFET process with two layers of gold metal. In order to lower capacitance, air bridges are used for isolating the metals. Because, currently, a GaAs MESFET process is still relatively expensive (the same price as a high performance silicon), the circuit fabricated is not very large or complicated. So, similar to MOSIS, multiple projects are built in one chip. The package has 44 pins, but packages with more pins are available (> 100 pins). The circuit fabricated has less than 44 pins, so several pins are shared by one signal for the purpose of reducing the inductance.

The fabrication in the QED process was also free of charge in exchange for the company’s interest. So, not much information about this process can be given here. The process files used for SPICE simulations are listed in Appendix A.2.

The SPICE simulation results in Fig. 70 for the transistor characteristics indicate that the drain current of a GaAs MESFET is comparable to the drain current of CMOS. Due to depletion-mode transistor used, the threshold voltage is less than zero (see Fig. 70a) and the saturation voltage, $V_{ds, sat}$, is relatively small. It also can be seen from Fig. 70b that the output impedance is very small, so a special technique should be employed for output impedance enhancement. Only N-channel transistors exist in this process.

VI.2. DESCRIPTIONS OF THE LAYOUT TOOLS

The layout tool used for the MOSIS chip is called Magic which was developed by the University of California at Berkeley [248]. The layouts for DABIC and QED were drawn in Allegro Microsystems, Inc. and TriQuint Semiconductor, Inc., so the tools used are company-proprietary. In this section, the only description given will be Magic.

Magic is an interactive system for drawing and modifying circuit layouts. With Magic, basic cells can be drawn by using a color graphic display and a mouse, and can be
Figure 70. SPICE simulation results for the drain current vs. gate-to-source voltage characteristic (a) and the drain current vs. drain-to-source voltage characteristic (b) of a GaAs NMESFET from TriQuint Semiconductor, Inc.
combined into a large structure hierarchically. Magic has built-in knowledge of design rule checking. So, during editing, Magic continuously checks for the design violations so that violations can be corrected before the whole circuit is laid out. Magic also knows about connectivity and transistors, and contains a built-in hierarchical circuit extractor.

Magic uses simplified design rules and circuit structures, so is easy to use and saves layout time. But, it costs some density of layout.

The commands of Magic can be entered by two different ways: typing the command starts with a semicolon or by using macros. The basic commands of Magic are paint, erase, copy, move, select, label, grid, view, zoom, undo, redo, and save. Magic also has many other commands including the straighten and plow which are very powerful ones for editing.

The layout data generated by Magic are saved in a file with .mag extension, and the data also can be extracted to an .ext file for post-layout performance verification (see Section VI.5 for details). For fabrication, a CIF file should be generated for making masks.

VI.3. DESIGN CONSIDERATIONS FOR LAYOUT

Before introducing the layouts, the design considerations should be discussed. Basically, the considerations can be sorted into four categories: process tolerance, parasitics, reliability, and design automation. As to process tolerance, there is little one can do. But for parasitics, many design issues are to be taken care of. Reliability, ideally, is not very important for a small circuit with several hundred transistors. Design automation, actually, is not a part of this thesis, but is considered briefly in this section.

VI.3.1. Parasitics

To reduce parasitics in a layout or to reduce their effect is very important for lay-
out design. In the following, the design considerations will be discussed from several different aspects.

**A. CMOS Latch-up.** CMOS latch-up is a common and important phenomenon in CMOS structures and is caused by parasitics. It cannot be observed from circuit design and, therefore, can only be avoided in layout design.

Figure 71a shows a cross-section of the layout of a P-well CMOS structure, building on a N-substrate and Fig. 71b shows its top view. On the left side of the figure, there is a PMOSFET containing the P-type drain/source diffusions, poly gate, and the aluminum interconnection. In the P-well, there is an NMOSFET which has a similar structure as the PMOSFET, except that N-type, instead of P-type, diffusions are used for the drain/source. In order to isolate these two devices, a thick silicon oxidation is used for a high threshold voltage.

Besides the above-mentioned real devices, there exist parasitic elements. One parasitic PNP lateral bipolar transistor is formed by P-diffusion, N-substrate, and P-well. Another parasitic NPN vertical transistor is built by N-substrate, P-well, and N-diffusion. These two transistors form a npnpn structure and are connected to each other as shown in Fig. 72 with \( R_W \) as the well resistance and \( R_S \) as the substrate resistance.

From Fig. 72, it can be derived that, if both of the base-emitter junctions of the bipolar devices are forward biased and if the loop gain satisfies the condition

\[
\beta_{PNP} \beta_{NPN} \geq 1
\]  

(VI.1)

with \( \beta_{PNP} \) and \( \beta_{NPN} \) being the current gains of the two transistors, the currents in both transistors will increase until they reach self-limit or until they result in the destruction of the chip or its bonding leads. This phenomenon is called CMOS latch-up.

There are several methods for preventing latch-up. One approach is to increase the distance between the drain/source diffusion of the P-channel transistor and the P-
Figure 7.1. A cross-section of the layout of a p-well CMOS structure (a) and its top view (b).

Figure 7.2. Basic latch loop formed by parasitic bipolar transistors and parasitic resistors in CMOS structure.
well. This reduces the value of $\beta_{PNP}$ so that the condition in Eq. (VI.1) becomes invalid. But it costs circuit density. The second approach is to reduce the resistances of $R_W$ and $R_S$ so that the forward base-emitter biases to the NPN transistor and to the PNP transistor are less than 0.6 V and, therefore, these transistors are off. The method to reduce these resistances is to place a P+ collar (called a guard ring) around the periphery of the P-well or an N+ guard ring surrounding the P-channel device.

An additional method to prevent latch-up is to use a different process so that a well can be isolated. The examples are to use CMOS/SOS instead of CMOS/bulk or to place an oxide trench deep enough to reach the low-resistivity substrate around the well. In both of these processes, both parasitic bipolar transistors are destroyed so that latch-up never happens. But, usually, these steps are expensive.

Considering density and costs, the method of guard ring is chosen to prevent CMOS latch-up for this thesis. Another reason for choosing the guard ring method is that it also helps to prevent the possibility of an N+ diffusion in the P-well from being too close to the edge of the well, which possibly creates a $V_{DD}$ to $V_{SS}$ short.

**B. Parasitic Capacitances and Crosstalk.** It can be seen from Fig. 71a that there is a parasitic capacitance between diffusion and substrate/well associated with the drain and source terminals of a device. In practice, this kind of capacitance cannot be avoided and the value of the capacitance depends on doping, applied voltage, and diffusion area. The first and second parameters cannot be changed in the layout design, but the third parameter can. The method to reduce the diffusion area is to share a diffusion by several different transistors. For example, the drain of the transistor M9 and the source of the transistor M11 in Fig. 45 are connected together, so that a P-diffusion can be shared by both transistors (the shaded area shown in Fig. 73). The method also can be applied to a single transistor by using multiple legs. In Fig. 74, a transistor with N $\mu$m width is laid out in one leg (Fig. 74a) and three legs (Fig. 74b), respectively. It can be easily calculated
with $D$ is the minimum width of diffusion, and the total diffusion area for the drain or source of the three-leg transistor is $((2/3)N\times D) \mu m$. So, the diffusion area is reduced by $1/3$, which results in reduced parasitic capacitances.

![The shared diffusion area](image)

**Figure 73.** Layout of two MOS transistors with the shared diffusion (shaded area).

Another kind of parasitic capacitance is wire capacitance which exists between poly/metal to substrate, poly2 to poly1, metal to poly, and metal2 to metal1. This kind of capacitance depends only on the area, so that the only way to reduce the capacitance is to use a narrow and short interconnection. In all filters designed to date, the current is not very large, so that a minimum wire width will be used. Also, the length of wire will be kept as short as possible.

If a capacitance exists between two signal lines, it will cause crosstalk, which means the signal in one line is being passed to another line in error. To reduce crosstalk, routing two lines in parallel should be avoided. Also, the space between two wires may need to be large for reducing the parasitic capacitance.

**C. Substrate Noise.** The substrate noise is caused by the bulk resistances and the
parasitic capacitances. When signal voltages change at device terminals of devices, they will generate the corresponding changes in the substrate through the above parasitics. In digital circuits, the substrate noise is very large because the signal voltages vary from $V_{SS}$ to $V_{DD}$. In analog circuits, since the voltages only swing in a linear range, the noise is much smaller. In a mixed digital/analog circuit layout, reducing the substrate noise transferred from digital circuits to analog circuits is a very interesting and important topic.

The most common method for reducing the substrate noise is to use substrate taps. In this method, P+ or N+ deep diffusions are plugged in the substrate to reduce the bulk resistances on the surface of the substrate. These taps are connected to a DC voltage so that the potential on the surface is near constant.
Another advantage of using substrate taps is to reduce the crosstalk, specifically for analog filters, among OTAs in a filter. It should be pointed out that guard rings are also substrate taps which help reducing the substrate noise and crosstalk.

Because there are no digital circuits in this project, the substrate noise is not a very important issue. But, considering that the OTAs may be used for any kind of large analog signal processing circuits or for a large analog filtering system designed by design automation in the future, substrate taps will be placed at as many places as possible.

VI.3.2. Process Tolerances

Process tolerances are the most important sources causing unpredicted parameters in the fabricated circuits. The range of process tolerances are usually 10% to 20%. Generally, process tolerances are caused by the misalignments of the masks, the variations of the time and temperature used in diffusions and oxidations, changes of the dose of chemical components, the variations of the energy for ion-implantation, and so on. Process tolerances are very complicated physical phenomena, so, in this section, the considerations will be focused only on the changes of transistor and capacitor sizes which are caused by process tolerances.

A. Variations of Transistor Sizes. In circuit design, changes in transistor sizes caused by fabrication tolerance are a big problem for designers, because the changes are very hard to predict. As a result, changes of transistor sizes cause current variations in the required paths, and, most important, cause mismatches between current mirrors, differential pairs, and so on. This not only distorts the performances of the designed circuits, but may even lead the circuit to malfunction. Thus, to reduce process tolerances on transistor sizes becomes very important in layout design. Unfortunately, currently, there are not many efficient methods available for this purpose. One method that can partially help is to reduce the size mismatch among transistors.
One method for reducing the size mismatch is to place the transistors in layout as close to each other as possible. It stands to the reason that the differences of process variations between two transistors close to each other are minimum. For example, in practice, doping distribution cannot be uniform, but its change is continuous so that the difference between doping concentrations at two points is an inverse function of the distance between these two points. If two transistors are placed close enough in layout, the differences of the doping concentrations can be ignored. Similarly, the temperature difference between two points is also an inverse function of the distance, so that its negative impact can be minimum if two transistors are placed close together.

Another method for reducing the size mismatch is to align the edges of devices in either a vertical or a horizontal direction. This can reduce the mismatch caused by mask misalignment caused by the differences of lateral diffusion, and by differences of doping distributions. Also, usually, the devices are placed with transistor width either vertical or horizontal direction, but not both direction. The purpose is to reduce the misalignment.

Another method for reducing the mismatch among transistors is to design one transistor by using multiple legs. Different from the considerations for parasitics, the leg used here should have an unit size. For instance, if 3 μm is used as a leg, a device with 3 μm can be built with one leg and a device with 15 μm can be designed with 5 legs. This kind of arrangement can reduce mismatch because the mismatch is caused by the ratio of the variations, not variations themselves. If we use the above example, the ratio of the variations is

\[
\frac{3 + \Delta W + 3 + \Delta W + 3 + \Delta W + 3 + \Delta W + 3 + \Delta W}{3 + \Delta W} = \frac{5(3 + \Delta W)}{3 + \Delta W} = 5 \quad (01.2)
\]

where \(\Delta W\) is the variation caused by process tolerances. Of course, in practice, the variation to each transistor is different, so the variations cannot be perfectly cancelled. But, the ratio of the variations can be minimized.
B. Changes of Capacitances. In IC processes, there are two kinds of capacitors: gate capacitors and layer capacitors. The gate capacitor is formed by poly-gate and channel. Because the thickness of gate oxide determines the threshold voltage of a device, the gate oxidation has the best quality compared to other oxidations, which results in accuracy in capacitance. Also, because the thickness of the gate oxidation is the thinnest one among all oxidations, the gate capacitance per unit area is the largest one, which leads to small area ($C_{\text{gate}} = 0.88 \, \text{fF/} \mu \text{m}^2$ for MOSIS). One of the disadvantages of the gate capacitor is that its capacitance depends on the applied gate voltage because the bulk resistance under the gate oxidation is relatively high and the charge distribution at the surface is not uniform and is dependent on bias voltage. One method to solve this problem is to use an additional process step to generate a highly doped channel under the gate oxidation (similar to a depletion-mode transistor). But this method depends on the fabrication process which is out of our control. Another method used by most circuit designers is to build a MOS transistor and connect drain and source together as one terminal of the capacitor and the gate as another. One of the problems of this configuration is that only a grounded capacitor can be designed, because the source/drain terminal should be connected to either $V_{DD}$ or $V_{SS}$, i.e., to AC ground. But generally, floating capacitors are needed in our designs. Another problem is that, although the bulk resistivity under the gate oxidation is reduced, it still results in capacitance variation with voltage. Because the bulk resistance is not zero, the charge distribution at the surface is not uniform except in devices operating in strong inversion. The last problem is that the additional area used for the drain/source diffusions reduces the advantage of small area.

The layer capacitances include any combinations of two layers. For instance, any layers to substrate or poly1, poly2 to metal1, and metal1 to metal2. Out of all layer capacitances, the capacitance built by poly2 over poly1 is the best choice (actually, poly2 is, generally, used for capacitance). First, this capacitance has the largest value per unit area...
among all layer capacitances; it is approximately half of the gate capacitance. For example, it can be seen from the MOSIS process file in Appendix A.1 that $C_{\text{poly}} = 0.455 \text{ fF/}\mu\text{m}^2$. Second, the quality of the oxidation between poly1 and poly2 is very good, therefore, this capacitance is stable. Third, this capacitance is totally independent of bias voltage because both poly layers are conductors. Fourth, this capacitance is floating because neither poly layers are connected to substrate which is connected to $V_{SS}$. So, currently, this kind of capacitance is widely used in integrated circuits.

Because the structure of the capacitance is simple, there are only two things to go wrong with the process tolerances: edge effects and variations in dielectric thickness. Edge effects include the misalignments in the edges of two plates of the capacitor and the edge field effect caused by the curved electric field lines (fringing) near the edge of the capacitor. In order to reduce edge effects, a basic and easy method is to use large area so that the tolerance caused by edge effects is relatively small. Another popular method is to make one plate (usually the top plate) smaller than the other (see Fig. 75a). Then, a small misalignment of the top plate will not cause changes in the capacitance area. Although this method cannot avoid the edge field effect (actually, no method can), it helps to account for this effect easily [158].

Different from preventing edge effects, the method to reduce variations in dielectric thickness requesting a small area because the dielectric thickness is a monotonic function of size. Another method to avoid the effects of variations in dielectric thickness is to use small unit capacitances. It is well known that, in many applications, an accurate capacitance is not required but an accurate capacitance ratio. For example, in a transconductance-C filter, the quality factor is determined by the ratio of two capacitances (see Chapter IV). So, in order to obtain an accurate ratio, small unit capacitance should be used.

Figure 75b shows a capacitor, $C_2$, built by one single plate, with capacitance
three times larger than that in Fig. 75a, C1. The capacitance ratio between C2 and C1 is not accurate. Fig. 75b also shows another capacitor, C3, built by three small capacitors, with the same capacitance as C2. The capacitance ratio between C3 and C1 can be very accurate. The reason is the same as the one used for explaining the transistor ratio of the transistors with several legs. But, it should be pointed out here, C3 takes more area than C2 because of the spaces among the small capacitors.

Figure. 75. Design considerations of typical integrated capacitors for the edge effects (a) and capacitance ratio-matching (b).
VI.3.3. Reliability

In layout design, most of the reliability issues are taken care of by the design rules which, as introduced in Section VI.2, are included in the tools which automatically and continuously check for the design rule violations. In the following, only a few considerations which cannot be checked by the tools will be introduced.

A. Electromigration. In integrated circuits, aluminum is often used as conductor for interconnection of components. In aluminum, the current density is limited by a phenomenon called electromigration (EM). EM is the movement of atoms with currents flowing and can be likened to wind erosion in sand. If significant electromigration occurs, the conductors can become open. So, in the layout design, the metal should be wide enough so that the required current is smaller than the allowed current or EM current. For aluminum, the EM current is around $I_{mA/\mu m^2}$.

For EM consideration, the current of concern only flows in one direction. So, EM occurs on most of the metal wires including the ones for power supplies, for current sources, for current mirrors, etc. The current used to check the EM requirement is the average current instead of the real current. Because the circuits designed here are relatively small, most currents do not cause EM problem. The only wires to be concerned about are power supply lines.

B. Self Heating. The phenomenon of self heating is very similar to electromigration, but the current flows bi-directionally, such as the output current of OTAs. Self heating will cause the metal to warm and even to open. The way to avoid self heating is to use wide metal lines.

Usually, there is less worry about self heating than EM, because the current used to check the self heating requirement is not the peak but the root-mean-square value of current. For the filters designed in thesis, a minimum line width can be used.
C. \(\alpha\) Particles. When an \(\alpha\) particle hits silicon, a large current can occur if the impact point has small capacitance. The reason is that an \(\alpha\) particle has enough energy to break the crystal bonds and generate electron-hole pairs. A freed electron and hole can break other bonds to generate more and more electron-hole pairs. The \(\alpha\) particle phenomenon will not damage the circuit immediately, but will reduce the lifetime of the silicon.

The probability of \(\alpha\) particle happening is determined by the chip area, doping concentration, node capacitance, and so on. For our design, because the circuit has fewer than 500 transistors, the chip area is relatively small so that \(\alpha\) particle issues can be ignored. Also, the node capacitances in the OTA are usually larger than the minimum capacitance required to prevent \(\alpha\) particle effect, say typically 0.07 \(pF\).

VI.3.4. Design Automation

This thesis is not concerned with design automation. But, as mentioned in the introduction, one advantage of transconductance-C systems is that they are good candidates for design automation [161, 166]. So, in order to make the designed OTAs available for the standard library of transconductance-C systems, the basic requirement for design automation is considered.

First consideration for design automation is the shape of OTA layout. If a system contains hundreds of OTAs, the shape is not very important. But, if they only contain several tens of OTAs, a good shape will save time for the placement and save chip area. Basically, an OTA with square shape is easy to place in systems and wastes less chip area than other kinds of shapes, because, generally, the die is square.

Another consideration for design automation is the pin arrangement. Because the connection between OTAs is, generally, in series, the input and output pins are placed on opposite sides. The control pins are better placed on the sides other than those for the
input and the output pins for easy routing and short connections. Power pins are usually placed on the top or bottom of the OTA so that connections can be as short as possible. Except for chip area constraints, in analog layout it should be avoided to place interconnections cross circuit components for reducing the parasitic capacitances.

The last consideration is to reduce parasitics. In design automation, OTAs are used as cells or a black boxes, so there is no special treatment for parasitics. Therefore, using short interconnections is important. If the circuit density permits, avoiding the use of overlapped layers helps reducing the parasitics.

VI.4. LAYOUT DESIGNS

In the last three sections, the fabrication process, layout graphic tools, and the design considerations were introduced; in this section, the layout will be drawn and be explained.

VI.4.1. Layout in MOSIS

The OTA based on the current addition technique and the elliptic filter introduced in Section IV.2 are drawn in the VLSI Computer-Aided Design Lab in the Department of Electrical Engineering at Portland State University, by using the Magic tool for the MOSIS process. As mentioned in Section VI.1.1, MOSIS has a scalable process. Therefore, the draw size is in a relative unit of lambda, not in the absolutely unit of μm. In a MOSIS process, for any layers the minimum spaces or feature sizes are 2 lambda wide. So, the unit lambda is converted to 1.5 μm for a 3 μm fabrication, to 1.0 μm for a 2 μm fabrication, to a 0.8 μm for a 1.6 μm fabrication, and to a 0.6 μm for a 1.2 μm fabrication.

In Magic, the layers are presented by rectangle boxes with shading inside and different colors for the box. In Fig. 76, all available layers for the MOSIS process are shown. With these layers, circuits can be drawn by a color graphic debugging tool, such as Magic. In the following sections, the descriptions for the layouts of the OTA and filter
### Layer Names and Colors

<table>
<thead>
<tr>
<th>Layer</th>
<th>CIF</th>
<th>CALMA</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>WELL</td>
<td>CWG</td>
<td>53</td>
<td></td>
</tr>
<tr>
<td>PWELL</td>
<td>CWP</td>
<td>41</td>
<td></td>
</tr>
<tr>
<td>NWELL</td>
<td>CWN</td>
<td>42</td>
<td></td>
</tr>
<tr>
<td>ACTIVE</td>
<td>CAA</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>SELECT</td>
<td>CSG</td>
<td>54</td>
<td></td>
</tr>
<tr>
<td>PSELECT</td>
<td>CSP</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td>NSELECT</td>
<td>CSN</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>POLY</td>
<td>CPG</td>
<td>46</td>
<td></td>
</tr>
<tr>
<td>CONT TO POLY</td>
<td>CCP</td>
<td>47</td>
<td></td>
</tr>
<tr>
<td>CONT TO ACT</td>
<td>CCA</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td>METAL1</td>
<td>CMF</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td>VIA</td>
<td>CVA</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td>METAL2</td>
<td>CMS</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>CONT TO ELEC</td>
<td>CCE</td>
<td>55</td>
<td></td>
</tr>
<tr>
<td>ELECTRODE</td>
<td>CEL</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>OVERGLASS</td>
<td>COG</td>
<td>52</td>
<td></td>
</tr>
</tbody>
</table>

Figure 76. Layer names and colors of MOSIS process.
A MOSIS OTA. Figure 77 shows the layout for the CMOS OTA in Fig. 45. For design automation, the shape of the OTA is designed near square (310 μm x 290 μm). The input pins are on the left and the output pins on the right side. The positive power supply $V_{DD}$ is at the top side of the OTA, and $V_{SS}$ is at the bottom. The control pins, $I_s$, $V_2$, and $-V_1$ are close to the bottom side and the control pins, $V_c$ and $V_1$, are close to the top.

It can be observed from Fig. 77 that P-channel transistors are placed together so that the number of N-wells used is minimized. The purpose is to save chip area because the space between the wells is the largest one in the design rules. In this design, only two N-wells are used. Guard rings are placed surrounding both N-wells to prevent CMOS latch-up. As mentioned in Section VI.3.1.B, the guard ring also helps to keep from crosstalk between N- and P-channel transistors. To reduce the substrate noise, substrate taps are placed underneath the negative power supply and near the bottom of the right side of the OTA. Also, the substrate taps in conjunction with the guard ring help to isolate the OTA from all sides except the part at the top. But, because two capacitors are placed on the top side, the space from the signal lines to the outside is too large for noise coupling.

In order to reduce the mismatch caused by process tolerances, the transistors M1 to M6 in Fig. 45 are placed together. Also, the transistors in the current mirrors, (M7 to M9), (M8 to M10), (M24 to M25), and (M27 to M20, M21, M22), are put very close to each other. The other adjacent transistors are (M13, M14), (M15, M16), and (M17, M18). There are two mismatched different pairs, (M1, M2) and (M5, M6) and one mismatched current source (M20/M22, M27). Because the area ratio between M27 to M20 (M22) is two, M27 is designed in four legs and M20 (M22) is in two legs. The purpose is to reduce the mismatch (a detailed explanation was given in Section VI.3.2.A). For M1 to M2 (M6
Figure 77. Layout of the CMOS OTA with current addition shown in Fig. 45.
to M5), because the ratio is as large as 80/3, it is impractical to use 3 μm for basic unit. In the layout, M1 (M6) is designed with four legs and M2 (M5) is in one leg.

Another purpose for using multiple legs is to reduce the diffusion area for minimizing the parasitic capacitances. So, although the ratios for the current mirrors, (M7, M9), (M8, M10), and (M24, M25), are one, multiple legs are also used. In addition, if possible, the transistors are laid out with shared diffusion area for the same purpose. For example, it can be seen from Fig. 77 that the diffusion nodes between two N-channel transistors in the output stage, (M13, M15) or (M14, M16), are shared by these two transistors.

There are two capacitors for the phase compensation discussed in Section II.3.3.A (see Fig. 25). The capacitors were made by the metal1 over poly1, because at the time of lay out the poly2 layer was not ready in MOSIS. After poly2 became available, they were not changed because their capacitances are small. As can be seen from the layout, the top plate is smaller than the bottom one to reduce the edge effects. Because the capacitance is not very large, the multiple-unit capacitor is not used for the reduction of variations in dielectric thickness.

Figure 77 shows that most wires do not overlap each other. This has the least parasitic capacitances, but costs chip area. Because the parasitics are more important than area in this design, the OTA is not compacted. Another gain obtained from this arrangement is that crosstalk is reduced.

For EM considerations, the metal width used for power supply lines is 8 μm. As mentioned in Section VI.3.3, self heating and α particles are of little concern.

B. Third-Order Elliptic Filter. Figure 78 shows the microphotograph of the overall MOSIS chip which includes two OTAs and two third-order elliptic filters. The OTA located at the left side of the center uses the current addition technique; its layout was discussed above. Another OTA, located at the right side of the center, is not used and discussed in this dissertation. Both filters have the same structure shown in Fig. 44 and
are built by the same OTA with the current addition. The difference between two filters is that one filter has a 20 MHz frequency response (on the top half of chip), and the other one is for 12 MHz (on the bottom half). Each one has seven OTAs (see Fig. 44): counting clockwisely, the first OTA starts from the left side and the last one ends at the right side. The OTAs are the same for both filters, but the circuit capacitors are different. The reason for such arrangement is that, as described in Section IV.2.2, the filter with 20 MHz frequency response is designed to reach the highest obtainable frequency response, so, it is risky. If the chip does not work for 20 MHz because the circuit capacitances are close to the parasitics, the 12 MHz filter still has chance to work because the circuit capacitors for this filter are larger. There is enough area and pins to permit this arrangement.

As mentioned earlier, the best way to connect the OTAs in a filter is in series. Because the chip die is nearly square, the seven OTAs cannot be placed on one line in the horizontal direction. Therefore, they are placed as shown in Fig. 78.

On the top level hierarchy, there are only OTAs as building blocks, interconnections, capacitors, and output buffers. The four-transistor output buffers are placed on the right side of the chip and close to the output OTA. The purpose of this is to shorten the connections from the OTAs to the buffers and from the buffers to the pins as much as possible. Generally, there is some work for arranging the routing between the OTAs due to their fixed positions. But much effort is needed for arranging the routing of the signals of circuits to the pins, since the connections between pins and signals are not unique.

The circuit capacitors of the filters are built by poly2 over poly1 to save area and to reduce the variations in dielectric thickness. The poly2 plate dimensions at least are 3 μm less than those of poly1 plate on each side to reduce edge effects. The quality factors $Q_{oi}$ determined by ratios of capacitors are not very large. Thus, no division into unit capacitors was used to save area. Note that the reason for the many empty spaces left in the layout is that poly2 was not available at the beginning of this layout design; so, the
Figure 78. Microphotograph of the third-order elliptic filter (see Fig. 44) fabricated by the MOSIS vendor.
capacitors were built by metal1 over poly. After poly2 was ready, the capacitors are re-laid out, but the area was not compacted because of the amount of work.

Although the circuit capacitors in Fig. 44 are the floating, grounded capacitors were used for all capacitors except $C_2$, because in a capacitor built in poly2 over poly1, there is a parasitic capacitor between poly1 and substrate. In practice, the best way to avoid this parasitic capacitor is to connect poly1 to substrate, which shorts the parasitic capacitor, but also results in a grounded capacitor. For the capacitor $C_2$, the floating capacitor must be used, so that a parasitic capacitor cannot be avoided. It can be seen from Fig. 44 that this parasitic capacitance will not affect the capacitance of $C_2$, but that of $C_1$ if poly1 of $C_2$ is connected to $C_1$, or the capacitance of $C_3$ otherwise. So, in the layout design, this parasitic capacitance should be subtracted from $C_1$ or $C_3$ as determined by connections. In this layout, $C_1$ is predistorted to account for the parasitic capacitance of $C_2$.

It can be seen from Figs. 44 and 45 that there are 11 pins for the OTA and 12 pins for the filter (one more for the bias voltage of the output buffers). Note that, the pin for $V_C$, which is used for adjusting the phase (see Fig. 25b), is not shown in Fig. 45. There are 7 pins for the OTA which is not used in this thesis, so a total of 42 pins for the complete chip. In MOSIS, either 40 or 64 pins are available. In order to reduce the cost, the 40-pin package is used with several shared pins. In the layout design, the pin $-V_1$ are shared by the both filters, and the pin $V_{SF}$ is shared by the unused and used OTAs. The pin assignments are listed Appendix B.1.

In CMOS technology, the pins connected to the gates of transistors should be protected to prevent accidental destruction of the gate oxide caused by electrostatic discharge (ESD). A typical pin ESD protection circuit is shown in Fig. 79a, which contains two reverse-biased diodes and a resistor. If a large voltage is applied to a pin, one of the diodes will break down depending on the polarity of the applied voltage. If the resis-
tance of resistor is large enough, the breakdown current is limited, therefore, the diode is not destroyed. The cross section of the layout for the protection circuit in Fig. 79a is shown in Fig. 79b.

![Diagram of CMOS ESD pin protection circuit](image)

Figure. 79. The equivalent CMOS ESD pin protection circuit built by P+ and N+ diffusions (a) and the cross-section of the layout (b).

In the MOSIS process, the pin ESD protection is provided automatically so, there is no need to do anything in the layout.

VI.4.2. Layout in DABIC

The OTA based on the current subtraction technique discussed in Section II.3 and the sixth-order Bessel filter introduced in Section IV.3 were laid out by Allegro Microsystems, Inc. The layout tools used are company-proprietary and the notations for layers are not clear. But, we believe all design considerations mentioned before are taken
care of in the layout design. In this section, only a brief discussions will be given below.

A. DABIC OTA. Figure 80 shows the layout for the CMOS OTA in Fig. 81. The reason the OTA circuit is redrawn here is that it has a small change: the bias voltage for the common-mode feedback circuit, $V_B$, in Fig. 19 is generated by a resistor and a transistor. From Fig. 81, it can be seen that, because cascode circuits are not used for the output stage (the transistors in DABIC have very large output impedance as shown in Fig. 69), the bias voltages to the output stages, $\pm V_1$ in Fig. 45, are not used. Also, the phase compensation circuits are not used in this OTA, and neither is the control voltage $V_C$ used in the MOSIS chip. So, there are only 7 pins used in this OTA.

Figure 80 shows that the unit leg used in this design is 50 $\mu$m long. So, the transistors in the bias generation circuits at the bottom of the layout have multiple legs; the transistor with the largest width has 8 legs. It is a little surprising that the input transistors for the mismatched differential pairs are not drawn in multiple legs because the width ratios of these transistors is only 4. All input transistors are laid out in one leg, shown in the middle of the layout. All P-channel transistors are placed on the top of the layout and the common-mode feedback circuit is on the right-top corner.

We were told this process is a double-well process, but only the P-well is visible in the layout. The reason may be that the N-well can be generated by using the P-well boundaries. There are a total of four wells. The resistor used for the bias circuit is on the right-top of the layout with an active diffusion area used underneath the resistor. It can be observed from Fig. 80 that the widths of the metal lines for power supplies are wide enough for preventing EM problems.

B. Sixth-Order Bessel Filter. Figure 82 shows the microphotograph of the overall chip from Allegro Microsystems, Inc., which contains a total of six circuits; three circuits, an OTA, a second-order biquad (see Fig. 38), and a sixth-order Bessel filter (see Fig. 49) are built by the current subtraction technique and are shown at the top half of the
Figure 80. Layout of the CMOS OTA with the current subtraction method shown in Fig. 6.14.
Figure 8.1 Circuit diagram for the OTA (current subtraction) with bias circuit to common-mode feedback circuit.
chip. Another three circuits, also an OTA, a biquad, and a sixth-order filter, are built by the current addition technique and are seen at the bottom half. Twelve OTAs in the sixth-order filter are placed in a straight line in the middle of the chip and the second-order biquad and the OTA are placed into the corner.

As mentioned in Section IV.4, the capacitors used in DABIC are built from several small capacitors with the unit capacitance of 0.25 pF. Also, all capacitors for the biquad are grounded. In the layout, many unit capacitors are placed between the OTAs. If a 0.75 pF-capacitance is needed, three unit capacitors are connected together in parallel. If 4 pF is needed, 16 units are connected. This is easy for the layout, but costs chip area. Fortunately, all circuits together have only 800 transistors, so that the area is small.

Because there are a total of 12 OTAs in the sixth-order filter, the current is relatively large. Thus, the metal lines with very wide widths are used for power supplies.

As mentioned above, there are seven pins for the OTA. For the sixth-order filter, there are also seven pins. For the second-order biquad, there are nine pins because it has both lowpass and bandpass output. For the three circuits with the current subtraction technique, all ground connections share one pin, and for the current addition, all ground nodes share one pin also. So, there are a total of 42 pins. With a 68 pin package, it is easy to make the pin arrangement; it is listed in Appendix B.2.

The pin ESD protection circuit of this chip is shown in Fig. 83 and its layout is shown in Fig. 84. Fig. 83 is very similar to Fig. 79. Note that the pin ESD protection circuits are only used for the signal pins, not for power supplies because they are never connected to the gate of transistors.

VI.4.3. Layout in QED

The GaAs OTA shown in Fig. 59 was laid out by TriQuint Semiconductor, Inc., neither the inductor discussed in Section V.2 nor the third-order elliptic filter mentioned
Figure 83. ESD pin protection circuit used in DABIC chip fabricated by Allegro Microsystems, Inc.

Figure 84. Layout of ESD pin protection circuit shown in Fig. 83.
in Section V.3 were fabricated because there was no chip area available. As mentioned in Section VI.1.3, fabrication in the QED process was free of charge. The QED process is a multi-project chip, so our circuit was placed only where a free space was available. Because the required chip area for the filter is relatively large, the chance to have a free space has not materialized yet.

Figure 85 shows the layout of the GaAs OTA. Since only N-channel transistors exist in this GaAs process, there is no latch-up problem. Transistors with multiple legs are not used because the sizes of all transistors in the OTA are relatively small (the maximum size is 40 μm). Also, because the chip area for the OTA is relatively large, the devices and interconnections are placed so far apart that parasitics are very small and the chances for crosstalk are minimum. The semi-insulating substrate reduces parasitic wiring capacitance.

It can be seen from Fig. 85 that the resistors used for the source degeneration in the GaAs OTA in Fig. 59, \( R_1 \) and \( R_2 \), are built by N-diffusions which is located in the middle of the layout (the transparent area). The capacitors used for the compensation of the excess phase shift, \( C_1 \) and \( C_2 \), are near the resistors. The capacitors for AC compensation, \( C_3 \) and \( C_4 \), are placed near the center; they are relatively large and seem to be built by metal over metal (MIM). The area used for either resistance or capacitance seems relatively small.

In the layout, power supply lines with the very wide widths are used for reducing the bonding inductances, not for electromigration because the current for the OTA is very small. Also, multiple pins are used for power supplies for the same purpose. The multiple pins are used for the outputs also because the current is relatively large, compared to the other pins.

It can be seen from Fig. 59 that there are a total of 9 pins for this OTA. But, in the layout, the bias generation circuit, which includes q23 and d24 to d33 in Fig. 59, was not
used. So, there are a total of 11 pins in the layout. The package has a total of 44 pins so several signals share multiple pins. The pin assignments are listed in Appendix B.3.

The microphotograph of the GaAs OTA is shown in Fig. 86.

VI.5. POST-LAYOUT PERFORMANCE VERIFICATIONS

After a layout is ready, the final step needed before fabrication is post-layout performance verification (PV) which includes the extraction of the parasitics and the simulation of the circuits with the layout parasitics in SPICE. This step also helps to check the circuit interconnections.

Because there is no layout information available for either DABIC or QED lay-
outs, no post-layout PV could be done at PSU. For the MOSIS, the layout was extracted from Magic to generate both the ota.ext and the filter.ext files. Then the ota.sim and filter.sim files are generated by using the comment "ext2sim". Finally, the SPICE files ota.spice and filter.spice files are generated by "sim2spice".

Unfortunately, when these programs are used to generate SPICE files, the programs cannot generate the capacitances properly, so the filter cannot be simulated. For the OTA, there is no problem because the capacitances are only used for phase compensation. The simulation results in Fig. 87 show that both the linearity and the linear input range of the OTA with current are very close to the simulation results in Fig. 47.

The results in Fig. 87 also show that the element connections of the OTA are correct, because they are close to those shown in Fig. 46. For the filter, the interconnec-
tions on the top level are easy to check manually. Because the capacitances are not available for simulations, the measured results will be used to check against the simulation results. In the next chapter, measurement and experimental results are presented.

Figure 87. SPICE simulation results based on the OTA circuit (Fig. 45) extracted from the layout shown in Fig. 77.
CHAPTER VII

EXPERIMENTAL RESULTS

With the layouts discussed in the last chapter, the circuits can be fabricated for measurements. All chips are measured in the Analog IC Test Lab in the Department of Electrical Engineering at Portland State University. In this chapter, the discussion of procedures for the measurement and the experimental results will be presented.

VII.1. CHIP FROM MOSIS

The overall chip from MOSIS includes two OTAs and two filters as discussed in Section VI.4.1. Only the OTA with the current addition method will be measured here. The filter with 20 MHz frequency response has been measured and works well, so the filter with 12 MHz is not measured. In the following, a detailed discussion will be given.

VII.1.1. CMOS OTA with Current Addition

For an OTA, the most interesting performance is linearity, and next is frequency response. The linearity was measured at DC, so it was easily obtained. The frequency response was measured at AC high frequencies and is very difficult to achieve. As discussed earlier, if the frequency response of an OTA is measured without an output buffer, the pin capacitance, in conjunction with the output resistance of the OTA or load resistance, will become the dominant pole because, usually, pin capacitance can be as large as a few pico-farads. On the other hand, if an output buffer is used, the properties of the OTA cannot be measured because the interesting output parameter of an OTA is current, rather than voltage. For example, with an output buffer, the AC output impedance of an
OTA cannot be measured. Also, the parameter $g_m$ of a voltage-controlled current source cannot be obtained. To date, there is still no reliable and direct method to test the AC performance of an OTA; rather one tests a filter for proving the performance of the OTA indirectly.

For testing the DC performance of the OTA, a power supply, a function generator, and a Tektronix's oscilloscope, were used. Fig. 88 shows the DC setup from which it can be seen that the outputs of the function generator are connected to the inputs of the OTA directly and the outputs of the OTA are sent to the inputs of the oscilloscope directly also. The value of the load resistor chosen is 1 kΩ. Under this environment, the DC transfer curve of the OTA obtained is shown in Fig. 89.

![Diagram](image)

**Figure. 88.** Experimental Setup for the DC Measurements.

It can be seen from Fig. 89 that this OTA has good linearity over a differential input range as large as ±6 V with ±5 V power supplies. Due to the lack of the necessary equipments, the slope of transfer curves, therefore, the transconductance cannot be obtained. But, it can be observed that there are ripples in the DC transfer curve, which correspond to the peaks in the transconductance in Fig. 21a.

It is also seen from Fig. 89 that the transconductance of this OTA is near 150 $\mu$S which is slight higher than that shown in Fig. 47 ($g_m = 140 \, \mu$S). The reason may be that due to process tolerances the sizes of the transistors are slightly larger than the required size. Thus, $g_m$ is a little larger and the input linear range is a little bit smaller than
The output voltage $V_{out}$ in Fig. 89 does not become flat for $|V_{out}| > 6$ V is because the differential-pair (M3, M4) is still on and provides current when the mismatched differential pairs (M1, M2) and (M5, M6) are turned already off.

Unfortunately, the acceptable linear tuning range of this OTA is only a factor of 1.3 which is relatively small and is caused by the transconductance of the transistors being relatively large. It can be seen from Fig. 20 that for a large $g_m$ the peaks of $g_{m1}$ and $g_{m2}$ are high. But, in order to avoid the nonlinear range of $g_{m1}$ ($g_{m2}$) from overlapping the linear range of $g_{m2}$ ($g_{m1}$), $g_{m1}$ ($g_{m2}$) should be as narrow as possible. Then, the nonlinearity in the center area (near the origin) is hard to recovered. Therefore, the linear tuning range is reduced.

Due to the differential outputs used, the DC offset is almost zero, as shown in Fig. 89. The DC current of this OTA is near 6 mA, which results in a total power dissipation of nearly 60 mW. The power dissipation is so high because the OTA has three different-
tial pairs instead of one as used in the most OTAs.

VII.1.2. CMOS Elliptic Filter

For the filter measurement, the most interesting parameter is the frequency response. Because the expected cutoff frequency of the designed filter is 20 MHz, the measured frequency range should be 100 MHz so that the notch and the stopband attenuation can be observed. This results in very difficult measurements. The difficulties in high-frequency measurements are the parasitics which associate with the pads, interconnection wires, and the leads of the components. For example, the leads of a resistor contain parasitic capacitance and parasitic inductance. So, when connecting all elements together, the interconnection wires, as well as the leads of the components, should be as short as possible.

The equipment used for this measurement are power supplies, the impedance/gain-phase analyzer HP4194A and the spectrum analyzer. The input signal is generated by the function generator which is contained in the HP4194A so that the references for phase and magnitude are easy to set up. The input signal is single-ended so that an off-chip transformer is used to generate differential inputs. The bandwidth of the transformer itself is 300 MHz. For the output side, because an on-chip buffer is used, an off-chip buffer is not necessary. The input to the phase and magnitude analyzer is single-ended, so another transformer is needed to generate a single-ended signal from the fully differential output. Because the transistor sizes in the output buffer are 300 µm, the output resistance is several hundred ohms. The input and output load resistors are two 100 Ω resistors in parallel so that they are matched to the 50 Ω cables which are used for the input and output connections. The experimental setup is shown in Fig. 90. Note that the bias current $I_b$ is generated by a variable resistor which is used to replace $I_b$ and is not shown in Fig. 90.
In order to reduce the noise from the power supplies, bypass capacitors are used for all the supplies and bias generators.

With the power supplies at ±5 V and the bias voltages as discussed in Section IV.2, the filter in Fig. 44 is measured and the experimental results for the frequency response are shown in Fig. 91. It can be seen that, with a careful setup, the frequency response curve is smooth in the frequency range: 1 MHz to 80 MHz. It can be seen from Fig. 91 that this filter works very well; the crossover frequency is near 20 MHz which is close to the design goal, the transmission zero is at 38 MHz with a notch depth greater than 60 dB, The stopband attenuation is 20 dB. It can be seen from Fig. 91 that the signal-to-noise ratio of this filter is near 70 dB. The noise output is obtained by connecting the input of the filter to ground. It should be pointed out that the low-frequency gain of this filter in Fig. 91 is near -50 dB because a 40 dB attenuator is used in the input. Some additional loss is caused by the transformers.

Besides the frequency response, the linearity of the filter, which affects the total harmonic distortion, is checked and the spectrum results are shown in Fig. 92. The fre-
Figure 91. Experimental results for the frequency response of the third-order elliptic filter.

frequency of the input signal is 3 MHz. Fig. 92a shows the input spectrum which indicates that the function generator is not perfect and contains distortions itself. Fig. 92b shows the output spectrum which indicates that distortion is added to the signal in the filter. By counting out the input spectrum from the output, the total harmonic distortion (THD) caused by the filter can be calculated as a function of the input voltage. Changing the input voltage and calculating the corresponding THD, the curve for THD vs the differential input voltage is shown in Fig. 93. It can be seen that for 2 V peak-to-peak input voltage THD is a respectable 1.05%.

Generally, the harmonic distortion of the low-order terms is larger than that of the high-order terms. It can be observed from Fig. 92, however, that the second harmonic at the output of the filter is less than the third harmonic. The reason is that, as mentioned earlier, in a fully-balanced structure the even-order harmonics, ideally, cancel each other, but not the odd-order terms.
Figure 92. Measured spectrum for the third-order elliptic filter; Input (a) and Output (b).
Figure 93. Total harmonic distortion vs the differential input peak-to-peak voltage.

As mentioned in the OTA measurement in Section VII.1.1, the linear tuning range of this OTA is not large, but only 30%. Also, it is discussed in connection with the phase tuning in Section II.3 that tuning also changes the phase. So, when the cutoff frequencies of the filter are changed by varying the bias current, the phase of the filter also changes. This situation can be seen easily during measurement from changing notch depths. Fortunately, the notch depth can be adjusted by phase tuning which was briefly discussed in Section II.3. In Fig. 94, the frequency responses with tuning are shown near the notch: both curves have a -45 dB notch depth. The frequency change at the positions of the notches is near 4 MHz. A larger tuning range can be reached with careful phase adjustment.

In Table IV, the experimental results of the filter are summarized.
Experimental results for the frequency responses of the third-order elliptic filter with tuning.

TABLE IV

TYPICAL PERFORMANCE CHARACTERISTICS OF THE CMOS ELLIPTIC FILTER

<table>
<thead>
<tr>
<th>Specification</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chip area</td>
<td>6.5 mm²</td>
</tr>
<tr>
<td>Power supply</td>
<td>±5 V</td>
</tr>
<tr>
<td>Power dissipation</td>
<td>~500 mW</td>
</tr>
<tr>
<td>Passband edge frequency</td>
<td>~20 MHz</td>
</tr>
<tr>
<td>Stopband rejection (f ≥ 35 MHz)</td>
<td>&gt; 20 dB</td>
</tr>
<tr>
<td>Notch position</td>
<td>38 MHz</td>
</tr>
<tr>
<td>Differential input swing at 0.5% THD</td>
<td>1.2 Vp-p</td>
</tr>
<tr>
<td>Dynamic range (signal/noise)</td>
<td>70 dB</td>
</tr>
<tr>
<td>Supply rejection at 3 MHz</td>
<td>42 dB</td>
</tr>
</tbody>
</table>
VII.2. CHIP FROM ALLEGRO MICROSYSTEMS, INC.

The overall chip fabricated by Allegro Microsystems, Inc., includes two OTAs, two second-order biquads, and two sixth-order Bessel filters. But, only the OTA and the sixth-order Bessel filter designed with the current subtraction method will be measured because the OTA and the filter fabricated via MOSIS has already proven the design with current addition technique. Also, if the sixth-order filter works well, it is not necessary to check the second-order biquad. This means the biquad is fabricated only as a debug block; it will be useful if the sixth-order filter does not work. As mentioned earlier, the extra circuits were built because there was free silicon area.

VII.2.1. CMOS OTA with Current Subtraction

As mentioned in Section VII.1.1, a most important parameter of an OTA is linearity, and the frequency response of an OTA is very hard to obtain. So, for this OTA, only linearity will be checked. The equipment as well as the setup are the same as used for the MOSIS OTA. The power supplies used in this OTA is ±2.5 V (see Section IV.3), and the load resistance chosen is 1 kΩ.

Fig. 95 shows the experimental result for the DC transfer curve of the OTA with the current subtraction technique. The X-axis is the differential input voltage and the Y-axis is the single-ended output current. So, the transconductance of the OTA $dl_{out}/dV_{in}$ is near 80 μS which is close to the simulation result, 70 μS. It can be seen from Fig. 95 that the linearity of this OTA is very good from -2 V to +2 V and is acceptable from -2.5 V to 2.5 V. This is consistent with the predictions made earlier that this OTA has a very good linearity over a relatively small range, and the linearity tends to become worse very quickly out of that range (see Fig. 23). Note that the differential linear input range for the OTA with current addition method has ±6 V over a ±5 V power supplies.

In order to further investigate the linearity, the total harmonic distortion (THD) of
Figure 95. Experimental results for the DC transfer curve of the OTA with current subtraction.

this OTA is also checked by using a spectrum analyzer (same as done for the filter discussed in Section VII.1.2). The curve of THD vs. the differential input voltage is shown in Fig. 96. It can be observed that THD is less than 1% for the input voltage $V_{p-p} < 1.3$ V.

From Fig. 95, it can be seen that there is a small DC offset which is near 25 mV for a 1 kΩ load. The offset exists because the output measured is single-ended, not fully-balanced, so the common-mode offset is not cancelled. In the application circuit, the fully-differential outputs will be taken, and the offset will be substantially cancelled.

VII.2.2. CMOS Bessel Filter

For the chip from Allegro Microsystems, Inc., the test fixture was provided by the company; it is good for reduction of high-frequency effects caused by parasitics. For the measurement of the AC frequency response, the equipment used for testing this filter is
Figure 96. Total harmonic distortion vs the differential input peak-to-peak voltage for the OTA with current subtraction.

the same as that used for the MOSIS filter. Because the operating frequency range of this filter is in the megahertz frequency, the special treatment for high frequencies, discussed in Section VII.1.2, must be considered here also.

The experimental setup for the measurement of this filter is the same as that shown in Fig. 90; This means transformers are used at the input and output terminals to convert the single-ended signal to differential form, or vice versa. Also, bypass capacitors are used for reducing the noise from all power lines including bias voltage sources, and a variable resistor is used to replace the current source $I_s$.

Unfortunately, the layout designer in Allegro Microsystems, Inc., made a small, but severe mistake in the layout; The P-channel transistors, instead of the N-channel transistors, were drawn for M4 and M5 in the output buffer shown in Fig. 50. So, P-channel active common-source amplifiers are built, instead of source followers. It is well known that the properties of the common-source and common-drain stages are totally
different. For example, the former has a large while the latter has a small output impedance. The former amplifies the input voltage but the latter does not. The former has a 180 degrees phase shift from the input to the output but the latter has a zero phase shift.

This mistake was not found during the post-layout performance verification done in Allegro Microsystems, Inc., but was found during measurement. Allegro Microsystems, Inc., could not fabricate another chip, because the DABIC II process was replaced by a new process. This forced us to measure this defective chip.

Fortunately, the large output impedance of the common-source stage does not matter because this stage follows a common-collector stage which is formed by one open-emitter NPN transistor (see Fig. 50). The gain of common-source stage must be reduced to near unity, otherwise, DC bias in the output stage will be destroyed. The way to control the gain is to reduce the gate-to-source voltage $V_{gs}$ by setting the positive power supply to near 1 V because the gate voltage (the output voltage of the filter) is near 0 V. Also, the value of the load resistor should be small to reduce the gain. For the phase shift, the gain of the output stage is controlled to be close to unity because a small gain produces a large bandwidth, since the product of gain and bandwidth is approximately constant. Then, if the frequency range of interest is much smaller than the bandwidth, the phase of the output voltage is relatively near constant, 180 degrees.

The method to control the gain of the output stage is to use a low voltage for $V_{DD}$ and a small resistance for $R_L$. The problem caused by the small $V_{DD}$ is that the OTA operates under a low power supply. Fortunately, as mentioned before, one advantage of this OTA is that it can work for the low power supplies because three differential pairs are connected in parallel. But, the linear input range of the OTA, obviously, is reduced due to the lower power supplies, the transconductance of the OTA is reduced due to the decrease of $V_{gs}$ of the transistors in the current sources and $V_{ds}$ of all transistors, and the linearity of the OTA may become worse due to the reduction of $V_{ds}$. Also, tuning may be
destroyed because the output voltage changes a little during tuning. But for even a very small change in the output voltage, the gain of the output stage will change a lot and the earlier condition for making the circuit work may be broken.

The above analysis shows that, by using a low voltage for $V_{DD}$, the output stage does work, but the OTAs have many problems. Fortunately, the reduced input range is not a problem for testing because an input signal with small amplitude can be used. The small transconductance causes a reduction in the bandwidth, but some results can be obtained. The linearity of the OTA, dose not change much because drain-to-source voltage is only a second-order effect. Either tuning will be abandoned or the measurements will be affected by distortion.

Note that both the second-order biquad and the sixth-order filter have the same output buffers, so the biquad cannot overcome this limitation.

By using a 1 V power supply, an 8 $\Omega$ load resistor, and a input signal of several tens of millivolts, the sixth-order filter was made to work. The experimental results in Fig. 97 show that the low-frequency gain of this filter is near unity, which means the gain of the output stage is near unity as expected. Note that for this filter measurement, the attenuator is not used for the input because the input signal has a small amplitude.

It can be seen from Fig. 97 that the cutoff frequency of this filter is near 3 MHz which is 1 MHz less than the expected goal and is believed to be caused by the smaller transconductance. The signal-to-noise ratio is more than 70 dB which is large enough.

For a Bessel filter, the most interesting parameter is the linear phase, so the phase of this filter is measured. The result in Fig. 98 shows that the phase at low-frequencies is 180 degrees, instead of zero degrees, because the output buffer is a common-source stage. The phase has good linearity until near 10 MHz which is close to the simulation results in Fig. 51. Fig. 98 also shows that the delay of the filter is near 111 ns which is higher than the expected goal, 85 ns. The reason for this is that the transformer used in
Fig. 97. Experimental results of the AC transfer function and noise performance for the sixth-order Bessel filter.

Fig. 90 has a linear phase shift (see Fig. 99) which results in a constant delay 16.7 ns. Considering there are two transformers used for the input and output, the delay caused by the filter itself should be \((111 - 16.7 \times 2) = 77.6\) ns which is close to the expected value.

The power consumption of this filter is approximately 30 mW which is much lower than that of the MOSIS filter. The reason for this is that the power supply for this filter is low.

This filter works acceptably well in spite of the layout error. The last parameter to be checked is tuning which we predicted will not work well. The experimental results for tuning are shown in Fig. 100. It can be seen that, during the tuning, not only the cutoff frequency changes, but also the low-frequency gain of the filter. This is caused by the gain of the output stage changing with tuning. So, in practice, tuning does not work for this filter. But, with understanding the errors in the output buffer, the results in Fig. 100 show that tuning the OTAs works well in principle.
Figure 98. Experimental results of the phase as function of frequency for the sixth-order Bessel filter.

Figure 99. Phase of the transformer as function of frequency.
Figure. 100. Experimental results for the frequency responses of the Bessel filter with tuning.

VII.3. GaAs OTA FROM TRIQUINT SEMICONDUCTOR, INC.

The chip from TriQuint Semiconductors, Inc., contains only one GaAs OTA, because the area we received is too small to include all circuits of interest. The chip area is $60 \times 60$ mil$^2$.

As usual, the most interesting parameter of an OTA is linearity. The equipment used for DC measurements is the same as that used before. The measurement conditions are: +10 V and +5 V for the two positive power supplies, -5 V for the negative power supply, and a 10 kΩ resistor for the load (see Fig. 59). A large load resistor was used here because the magnitude of the input signal can only go to 0.5 V and more gain, $g_mR_L$, is needed to obtain repeatable measurements.

With the above setup, the DC transfer curve for single-ended input and single-ended output is shown in Fig. 101. It can be seen that the linearity of this OTA is good.
over -0.4 V to 0.4 V, which is consistent with the SPICE simulation results in Fig. 56 where the input is differential. Note that the differential input range is double the single-ended input range. When the input signal is 0.2 volt, the output voltage is near 0.4 V. Considering the load resistor is 10 kΩ, the transconductance of the OTA can be calculated as approximately 200 μS, which is less than the design goal of 267 μS (see Fig. 56). This is caused by two possibilities: process tolerances and low bias current $I_s$.

![Graph](image.png)

**Figure 101.** Experimental results for the DC transfer curve of the GaAs OTA.

The DC characteristics of the GaAs OTA are acceptable.

Because there is no GaAs filter in this chip, all other parameters on this OTA should be measured on the OTA itself. For example, the output impedance and the cutoff frequency response, which are both AC parameters. As mentioned in Section VII.1.2, to measure the AC performances at more than several tens of megahertz is very hard, even more so, as this OTA operates in the gigahertz frequency range. So, for the small-signal output impedance, only the low-frequency response will be measured.
The technique used to measure the low-frequency output impedance measures the output voltages for two different load resistors at low frequency. For example, $V_{o1}$ is the output voltage corresponding to the load resistor $R_{L1}$, and $V_{o2}$ for $R_{L2}$ (see Fig. 102). Then, the output resistor $R_o$ can be calculated by the following equation:

$$R_o = \frac{R_{L1} K - R_{L2}}{1 - K}$$

with

$$K = \frac{V_{o1} R_{L2}}{V_{o2} R_{L1}}$$

![Figure 102. Small-signal model of an OTA with the load resistor $R_L$.](image)

Table V shows the measured results at 1 MHz for calculating the output impedance.

**TABLE V**

<table>
<thead>
<tr>
<th>$R_L$</th>
<th>$V_{in}$</th>
<th>$V_o$</th>
<th>$A_v$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 kΩ</td>
<td>0.2 V</td>
<td>0.38 V</td>
<td>1.9</td>
</tr>
<tr>
<td>1 kΩ</td>
<td>0.2 V</td>
<td>0.04 V</td>
<td>0.2</td>
</tr>
</tbody>
</table>

From Table V, the low-frequency output impedance can be calculated as 170 kΩ. This value is a little bit larger than the SPICE simulation result, 160 kΩ. The reason for this is that, when the bias current is small, the transconductance will be smaller and the output
impedance will be larger.

For testing the frequency response, we received an engineering test fixture, specially for our package, from TriQuint Semiconductor, Inc. But, unfortunately, the available gain-phase analyzer, HP4194A, only can reach 100 MHz, and the oscilloscope, Tektronix 2247A, also can only reach 100 MHz. Without high-frequency equipment, the measurement is limited to 100 MHz.

Because an on-chip buffer cannot be used in an OTA, the measurement, as discussed earlier, is very difficult. Because of the large pin capacitance, a small resistance should be used to reduce the RC time constant of the load. But, for the high output impedance of the OTA, a large resistance is needed for matching; If not matched, the very large capacitance of a 50 Ω cable (usually, more than 100 pF) will load the OTA. So, in this test, the experimental setup in Fig. 90 is not used.

Fortunately, the pin capacitance with the test fixture from TriQuint Semiconductor, Inc., is near 1 ~ 2 pF. If a 1 kΩ load resistor is used, the frequency of the parasitic pole caused by the pin capacitance and the load resistance is still over 100 MHz. So, a 1 kΩ load resistor is used. For the interconnection, a X100 Tektronix high-frequency probe with 10 MΩ resistance and 0.7 pF capacitance is used. Of course, the probe is connected to the high-impedance connector on the HP4194A. Both the input and the output are single-ended and the bypass capacitors are used for the power supplies.

Figure 103 shows the experimental result measured from the above setup. Considering the x100 probe used, the gain of the OTA with a 1 kΩ load is -12.5 dB which indicates that the transconductance of this OTA is 243 μS, close to the expected value. The gain is almost flat from 0.1 MHz to 100 MHz with a only 0.5 dB drop which is believed to be caused by the the pole generated by the pin capacitance and the load resistor. So, Fig. 103 shows this OTA has a bandwidth more than 100 MHz.
Because the cutoff frequency cannot be obtained, the tuning capability cannot be found in an AC measurement. So, a DC measurement is done by checking the output voltage. When the digital control voltage $V_c$ in Fig. 59 is connected to ground and the analog control voltage changes from -2 V to 0 V, the output voltage changes by a factor 1.3. When $V_a$ is set to be 0 V and $V_c$ changes from -5 V to 0 V, the output voltage changes by a factor 1.4. So, the tuning looks fine. But, it should be pointed out here that the tuning measured is not continuous and may be out of the linear input range. Therefore, only a general picture about the tuning is given and a true test of tuning should be obtained from AC measurements.

The power dissipation of this OTA is around 15 mW.

VII.4. CONCLUSIONS

In this chapter, the CMOS OTAs built by the current addition and the current sub-
traction techniques were measured and the experimental results show that they work fine. This proves the designs are correct and the parasitics are controlled in an acceptable range. The corresponding filters were measured; one works very well and the other one works also, but with some special treatment for overcoming the layout errors. Finally, the GaAs OTA was tested with much effort for AC measurements. It also works reasonably well. So, all designs described from Chapter II to Chapter V have been verified by experimental results.
CHAPTER VIII

CONCLUSIONS

In this thesis, the design strategies for high-frequency OTA-based analog signal processing (ASP) systems have been presented. A number of filters, as examples of typical ASP systems, have been designed and evaluated to show these strategies. The advantages of these systems, such as easy design with simple design procedures and excellent for high-frequency performance, have been discussed and verified. New techniques for the design of the basic building block, the operational transconductance amplifier (OTA), have been proposed with emphasis on linearity, linear input range, frequency response, and common-mode feedback. New design techniques for the filtering system operating at a frequency range of several tens, even several hundreds, of megahertz have also been shown, with special attention paid to considerations for the high-frequency effects of parasitics. All circuits have been fabricated and evaluated experimentally.

For the design consideration of linearity and linear input range of an OTA, several new techniques were proposed in this thesis. One technique connects two CMOS cross-coupled pairs by DC voltage shifting stages [262] so that the non-linearity of the MOSFETs can be canceled. The OTA designed by this technique has very good linearity over a relatively small input range. In order to increase the input range, the P-channel transistors used in CMOS cross-coupled pairs are replaced by N-channel transistors, so that the resulting OTA has very good linearity over a large input range. The only problem is that its common-mode rejection ratio to unsymmetrical inputs is relatively poor. For improving common-mode rejection and power consumption, another new technique, called current addition, was proposed, which is based on a multiple-differential-pairs
structure. With this technique, the new OTA has good linearity over a large input range and high common-mode rejection ratio, but its performance seems relatively sensitive to process tolerances. A way to reduce this sensitivity is to use the current subtraction technique, but the linear input range is reduced. Generally speaking, different OTAs have different advantages, so it is necessary to select appropriate OTAs for different applications.

For the design consideration of common-mode rejection and power supply rejection of an OTA, the fully-balanced circuit structure is used. In addition, this structure helps to reduce the DC offset, to improve harmonic distortion, to enlarge the dynamic output range, and to minimize noise. With this kind of structure, common-mode feedback (CMF) circuits must be employed to stabilize the output DC voltage. Currently, most existing CMF circuits for op-amps cannot be directly used for OTAs because of the interaction between the operation of CMF circuits and OTA tuning. In order to reduce the impact of this problem, three new CMF circuits were designed: One has very simple circuitry with an acceptable DC offset, one has a zero DC offset with an acceptable reduction in the bandwidth, and the last one has a good performance but is only suitable for small output swing. Two of these circuits were used in the OTAs and filters designed in this thesis and their performance verified.

For consideration of frequency response, a new process, instead of AC compensation technique, was used in this thesis. Since an OTA is a current-mode amplifier, its bandwidth is much larger than that of an op-amp. Therefore, an AC compensation technique does not help much. In this thesis, GaAs MESFET technology was chosen for designing the first reported GaAs OTA.

In order to improve the low output impedance of GaAs MESFETs, a new technique for output impedance enhancement was presented, which is based on the self-bootstrapping technique with some modification for reducing the large output capacitance. In order to design the OTA with only N-channel GaAs MESFET, a new tuning
method was proposed, which is totally different from the traditional method for tuning in CMOS OTAs. Finally, in order to overcome the problem of small input swing of a MESFET, a digital tuning scheme was introduced. With these approaches, a 7-GHz OTA with linearity, output impedance, and tuning performance comparable to those of a CMOS OTA was designed.

With the new OTAs, filters were designed as applications and test vehicles of the OTAs. Because the design procedures for transconductance-C filter, cascade and LC ladder approaches, are well known, the design consideration in this thesis is focussed on reducing the high-frequency effects of parasitics. A new design method was proposed for this purpose, so that the designed filters have the highest obtainable frequency response.

Using CMOS OTAs built by the techniques of current addition and current subtraction, two high-frequency filters were designed: A 20-MHz third-order elliptic filter was simulated as an LC ladder structure in 2-μm CMOS technology, and a 4-MHz sixth-order cascade Bessel filter was designed in 3.5-μm CMOS technology. With the GaAs OTA, a 1-GHz GaAs inductor was designed using a new design procedure for achieving the highest obtainable frequency response. Also, a GaAs third-order elliptic filter with 200-MHz bandwidth was designed to show the performance of the transconductance-C system at very high frequencies.

In order to verify the designs, the OTAs and filters were laid out with design emphasis on three categories: parasitics, process tolerances, and reliability. The methods used in the layout design for reducing the parasitics and process tolerances and for increasing the reliability were presented. The circuits were fabricated through MOSIS, Allegro Microsystems, Inc., and TriQuint Semiconductor, Inc.

All chips were measured with an experimental setup built for reducing high-frequency effects caused by wires and contacts. The OTA with the current addition technique and the filter built with these OTAs, which were laid out at PSU and fabricated via
MOSIS, worked very well and the experimental results are very close to the simulation results. This means that the parasitics in these circuits are controlled in an acceptable range. The CMOS OTA with current subtraction, laid out and fabricated by Allegro Microsystems, Inc., also worked well, but, the corresponding sixth-order Bessel filter did not work under normal conditions because of a layout error. With a special method proposed to overcome the layout error, the obtained experimental results were not very accurate, but were good enough to verify the filter design. For the GaAs OTA, the DC performance was easy to test by experiment. But, due to the lack of a GaAs filter built on a chip, the AC performance could not be obtained. Although a very good AC test system was set up, the lack of high-frequency equipment became a bottleneck. Consequently, the AC performance of the GaAs OTA was only verified up to 100 MHz.

The circuit with the most complete results is the OTA built in the current addition technique. This OTA is the most robust and, currently, is used by the analog design group at PSU as a basic building block for design automation of analog filters [166].

For future work, a further step is to fabricate the 200 MHz GaAs filter and to find appropriate equipment for measurement. The results will be very interesting and attractive for future research in the field of high-frequency analog system designs.

Another interesting task for further research is to design an OTA with a simpler structure, which is the only design consideration not addressed in this thesis. Basically, improvements in linearity will increase the complexity of the OTA structure, but there must be a tradeoff point between linearity and complexity; Building an OTA with acceptable linearity in a simple structure is very important and interesting task.

The design of a system with low power dissipation is a very important task, and, as mentioned earlier, the OTA with multiple differential pairs is a good candidate. So, to simulate the OTA with multiple differential pairs for low power supplies and to find out the lower-limit of power supplies for this OTA is of interest.
To date, OTAs and OTA-based systems have been designed in CMOS, Bipolar, BiCMOS, and GaAs technologies. Because the different technologies have distinct advantages, such as current drive, parasitics, process tolerances, and cost, to find out which technology is the most suitable for the design of OTAs for given applications is a very interesting topic for further research.
REFERENCES


222


April 1990


APPENDIX A

PROCESS FILES
A.1 MOSIS PARAMETRIC TEST RESULTS

RUN: N11B / BARBARA_ALLEN
TECHNOLOGY: SCNA
VENDOR: ORBIT
FEATURE SIZE: 2.0μm

A.1.1. INTRODUCTION

This report contains the lot average results obtained by MOSIS from measurements of the MOSIS test structures on the selected wafers of this fabrication lot. The SPICE LEVEL 2 parameters obtained from similar measurements on these wafers are also attached.

A.1.2. TRANSISTOR PARAMETERS:

<table>
<thead>
<tr>
<th>PARAMETERS:</th>
<th>W/L</th>
<th>N-CHANNEL</th>
<th>P-CHANNEL</th>
<th>UNITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vth (Vds=.05V)</td>
<td>3/2</td>
<td>.886</td>
<td>-.959</td>
<td>V</td>
</tr>
<tr>
<td>Vth (Vds=.05V)</td>
<td>18/2</td>
<td>.816</td>
<td>-.941</td>
<td>V</td>
</tr>
<tr>
<td>Idss (Vgd=.5V)</td>
<td>18/2</td>
<td>2683.0</td>
<td>-1334.0</td>
<td>μA</td>
</tr>
<tr>
<td>Vpt (Id=1.0μA)</td>
<td>18/2</td>
<td>15.02</td>
<td>-15.55</td>
<td>V</td>
</tr>
<tr>
<td>Vth (Vds=.05V)</td>
<td>50/50</td>
<td>.836</td>
<td>-.911</td>
<td>V</td>
</tr>
<tr>
<td>Vbkd (Uj=1.0μA)</td>
<td>50/50</td>
<td>15.0</td>
<td>-16.5</td>
<td>V</td>
</tr>
<tr>
<td>Kp (Uo*Co/2)</td>
<td>50/50</td>
<td>25.1</td>
<td>10.56</td>
<td>μA/V²</td>
</tr>
<tr>
<td>Gamma</td>
<td>50/50</td>
<td>.186</td>
<td>.652</td>
<td>V₁/²</td>
</tr>
<tr>
<td>Delta Length</td>
<td></td>
<td>.325</td>
<td>.502</td>
<td>μm</td>
</tr>
<tr>
<td>Delta Width</td>
<td></td>
<td>.178</td>
<td>.331</td>
<td>μm</td>
</tr>
</tbody>
</table>

COMMENTS: These parameters seem normal.
A.1.3. FIELD OXIDE TRANSISTOR PARAMETERS:

<table>
<thead>
<tr>
<th>TRANSISTOR PARAMETERS:</th>
<th>GATE+N</th>
<th>SOURCE/DRAIN ACTIVE</th>
<th>SOURCE/DRAIN P+ ACTIVE</th>
<th>UNITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vth (Vbs=0,I=1μA)</td>
<td>Poly</td>
<td>17.3</td>
<td>-14.1</td>
<td>V</td>
</tr>
<tr>
<td>Vth (Vbs=0,I=1μA)</td>
<td>Metal1</td>
<td>20.4</td>
<td>-23.1</td>
<td>V</td>
</tr>
<tr>
<td>Vth (Vbs=0,I=1μA)</td>
<td>Metal2</td>
<td>20.5</td>
<td>-23.2</td>
<td>V</td>
</tr>
</tbody>
</table>

COMMENTS: These parameters seem normal.

A.1.4. CIRCUIT PARAMETERS:

<table>
<thead>
<tr>
<th>CIRCUIT PARAMETERS:</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vinv, K = 1</td>
<td>2.10 V</td>
</tr>
<tr>
<td>Vinv, K = 1.5</td>
<td>2.30 V</td>
</tr>
<tr>
<td>Vlow, K = 2.0</td>
<td>0.00 V</td>
</tr>
<tr>
<td>Vhigh, K = 2.0</td>
<td>4.99 V</td>
</tr>
<tr>
<td>Vinv, K = 2.0</td>
<td>2.45 V</td>
</tr>
<tr>
<td>Gain, K = 2.0</td>
<td>-11.76</td>
</tr>
<tr>
<td>Ring Oscillator Frequency</td>
<td>31.48 MHz (31 stages @ 5.0V)</td>
</tr>
</tbody>
</table>

COMMENTS: The ring oscillator frequency is typical.
A.1.5. PROCESS PARAMETERS:

<table>
<thead>
<tr>
<th>PROCESS PARAMETERS:</th>
<th>POLY</th>
<th>N POLY</th>
<th>P POLY</th>
<th>N DIFF</th>
<th>P DIFF</th>
<th>METAL 1</th>
<th>METAL 2</th>
<th>POLY 2</th>
<th>UNITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sheet Resistance</td>
<td>21.0</td>
<td>20.9</td>
<td>29.5</td>
<td>64.9</td>
<td>.048</td>
<td>.025</td>
<td>19.7</td>
<td>Ohm/sq</td>
<td></td>
</tr>
<tr>
<td>Width Variation</td>
<td>-.277</td>
<td>-.260</td>
<td>.293</td>
<td>.042</td>
<td>-.162</td>
<td>.111</td>
<td>-.178</td>
<td>μm</td>
<td></td>
</tr>
<tr>
<td>Contact Resist.</td>
<td>8.68</td>
<td>8.84</td>
<td>20.15</td>
<td>125.6</td>
<td>----</td>
<td>.031</td>
<td>8.91</td>
<td>Ohms</td>
<td></td>
</tr>
<tr>
<td>Gate Oxide Tox</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>383</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>Angst.</td>
<td></td>
</tr>
</tbody>
</table>

COMMENTS: These parameters seem normal.

A.1.6. CAPACITANCE PARAMETERS:

<table>
<thead>
<tr>
<th>CAPACITANCE PARAMETERS:</th>
<th>POLY</th>
<th>N DIFF</th>
<th>P DIFF</th>
<th>METAL 1</th>
<th>METAL 2</th>
<th>POLY 2</th>
<th>UNITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area Cap (Layer to subs)</td>
<td>.054</td>
<td>.109</td>
<td>.323</td>
<td>.027</td>
<td>.016</td>
<td>----</td>
<td>fF/μm²</td>
</tr>
<tr>
<td>Area Cap (Layer to Poly)</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>.041</td>
<td>.021</td>
<td>.455</td>
<td>fF/μm²</td>
</tr>
<tr>
<td>Area Cap (Layer to Metal1)</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>.036</td>
<td>.041</td>
<td>fF/μm²</td>
</tr>
<tr>
<td>Fringe Cap (Layer to subs)</td>
<td>----</td>
<td>.450</td>
<td>.330</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>fF/μm</td>
</tr>
</tbody>
</table>

COMMENTS: These parameters seem normal.

A.1.7. N11B SPICE LEVEL 2 PARAMETERS

MODEL CMOSN NMOS LEVEL=2 LD=0.224977U TOX=391.000000E-10
+ NSUB=9.632380E+14 VTO=0.812774 KP=5.321000E-05 GAMMA=0.19992
+ PHI=0.6 U0=603 UEXP=7.920621E-02 UCIR=5589.43
+ DELTA=1.94145 VMAX=61431.5 XJ=0.250000U LAMBDA=1.968774E-02
+ NFS=9.880855E+11 NEFF=1 NSS=1.000000E+10 TPG=1.000000
+ RSH=29.620000 CGDO=2.980350E-10 CGSO=2.980350E-10 CGBO=7.313744E-10
+ CJ=9.840000E-05 MJ=0.769300 CJSW=4.455000E-10 MJSW=0.290620 PB=0.800000
* Weff = Wdrawn - Delta_W, and the suggested Delta_W is 0.27 um

MODEL CMOSP PMOS LEVEL=2 LD=0.224676U TOX=391.000000E-10
+ NSUB=1.052030E+16 VTO=-0.969807 KP=1.996000E-05 GAMMA=0.6691
+ PHI=0.6 UO=226 UEXP=0.256135 UCRIT=61656.4
+ DELTA=0.517267 VMAX=1000000 XJ=0.250000U LAMBDA=5.203973E-02
+ NFS=1.000000E+11 NEFF=1.0011 NSS=1.000000E+10 TPG=-1.000000
+ RSH=66.240000 CGDO=2.976363E-10 CGSO=2.976363E-10 CGBO=7.957415E-10
+ CJ=3.252000E-04 MJ=0.566330 CJSW=3.172900E-10 MJSW=0.314400 PB=0.800000
* Weff = Wdrawn - Delta_W, and the suggested Delta_W is 0.53 um

A.2. THE GaAs QED PROCESS PARAMETERS

model qed2 subckt: nodes=(drain,gate,source) krs=1374 krd=1374 areads=0.64
  areadg=0.36 w n=1 gama=0.058 delta=0.56 coef=0.5
fctl=168/(w n)
delt=delta*fctl
area1=w n
dgsa=areads*(w+1)*n
dgda=areadg*(w+1)*n
rd drain idsp r: r=krs/(w n)
rs 3 source r: r=krd/(w n)
cds drain source c: c=0.2f w w*0.1f
dgs gate 3 d: area=dgsa is=2.63e-15 n=1.13 cjo=2.2e-15 pb=0.4 m=0.5 fc=0
dgd gate idsn d: area=dgda is=2.63e-15 n=1.13 cjo=2.2e-15 pb=0.4 m=0.5 fc=0
evgs 1 3 gate 3 2 3 vcvs: p1=coef p2=coef*gamma
cc1 10 0 ids p ids 2 cccs: p4=1
r10 10 0 r: r=1
vc1 3 ids 2 3 10 0 vccs: p4=frac
j1 2 1 3 jfet: area=area1 lambda=0 vto=-0.35 beta=4.72e-4 is=1e-30
endm qed2
APPENDIX B

PIN ASSIGNMENTS OF CHIPS
### B.1. PIN ASSIGNMENTS FOR MOSIS CHIP

<table>
<thead>
<tr>
<th>Pin</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-2</td>
<td>Outputs of the unused OTA</td>
</tr>
<tr>
<td>3</td>
<td>Bias voltage to the buffer of the filter with 20 MHz</td>
</tr>
<tr>
<td>4</td>
<td>$V_{DD}$ of the unused OTA</td>
</tr>
<tr>
<td>5-6</td>
<td>Outputs of the filter with 20 MHz</td>
</tr>
<tr>
<td>7</td>
<td>$V_1^{(1)}$ of the filter with 20 MHz</td>
</tr>
<tr>
<td>8</td>
<td>$V_C^{(2)}$ of the filter with 20 MHz</td>
</tr>
<tr>
<td>9</td>
<td>$V_{DD}$ of the filter with 20 MHz</td>
</tr>
<tr>
<td>10</td>
<td>Bias current of the filter with 20 MHz</td>
</tr>
<tr>
<td>11</td>
<td>$V_2^{(3)}$ of the filter with 20 MHz</td>
</tr>
<tr>
<td>12</td>
<td>$V_{SS}$ of the filter with 20 MHz</td>
</tr>
<tr>
<td>13-14</td>
<td>Inputs of the filter with 20 MHz</td>
</tr>
<tr>
<td>15</td>
<td>$V_1$ of the OTA</td>
</tr>
<tr>
<td>16</td>
<td>$V_{DD}$ of the OTA</td>
</tr>
<tr>
<td>17</td>
<td>$V_C$ of the OTA</td>
</tr>
<tr>
<td>18-19</td>
<td>Inputs of the OTA</td>
</tr>
<tr>
<td>20</td>
<td>Bias current of the OTA</td>
</tr>
<tr>
<td>21</td>
<td>$-V_1^{(1)}$ of the OTA</td>
</tr>
<tr>
<td>22</td>
<td>$V_2$ of the OTA</td>
</tr>
<tr>
<td>23-24</td>
<td>Outputs of the OTA</td>
</tr>
<tr>
<td>25-26</td>
<td>Inputs of the filter with 12 MHz</td>
</tr>
<tr>
<td>27</td>
<td>Bias current of the filter with 12 MHz</td>
</tr>
<tr>
<td>Pin</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>--------------------------------------------------</td>
</tr>
<tr>
<td>1 - 7</td>
<td>unused</td>
</tr>
<tr>
<td>8 - 9</td>
<td>Outputs of sixth-order filter with subtraction</td>
</tr>
<tr>
<td>10</td>
<td>Grounds of all circuits with subtraction</td>
</tr>
<tr>
<td>11 - 12</td>
<td>Bandpass outputs of second-order biquad with subtraction</td>
</tr>
<tr>
<td>13</td>
<td>$V_{CC}$ of second-order biquad with subtraction</td>
</tr>
<tr>
<td>14</td>
<td>Bias current of second-order biquad with subtraction</td>
</tr>
<tr>
<td>15 - 16</td>
<td>Inputs of second-order biquad with subtraction</td>
</tr>
</tbody>
</table>

Note

1. $V_1$ and $-V_1$ are the bias voltage to the output stage of the OTA.
2. $V_C$ is the control voltage to phase compensation in the OTA.
3. $V_2$ is the bias voltage to the common-mode feedback circuit of the OTA.

B.2. PIN ASSIGNMENTS FOR DABIC CHIP
17 - 18  Lowpass outputs of second-order biquad with subtraction
19  \( V_{CC} \) of OTA with subtraction
20, 24  Outputs of OTA with subtraction
21  Bias current of OTA with subtraction
22 - 23  Inputs of OTA with subtraction
25, 27  Inputs of sixth-order filter with subtraction
26  Bias current of sixth-order filter with subtraction
28  \( V_{CC} \) of sixth-order filter with subtraction
29 - 41  unused
42  \( V_{CC} \) of sixth-order filter with addition
43, 45  Inputs of sixth-order filter with addition
44  Bias current of sixth-order filter with addition
46, 50  Outputs of OTA with addition
47 - 48  Inputs of OTA with addition
49  Bias current of OTA with addition
51  \( V_{CC} \) of OTA with addition
52 - 53  Lowpass outputs of second-order biquad with addition
54 - 55  Inputs of second-order biquad with addition
56  Bias current of second-order biquad with addition
57  \( V_{CC} \) of second-order biquad with addition
58 - 59  Bandpass outputs of second-order biquad with addition
60  Grounds of all circuits with addition
61 - 62  Outputs of sixth-order filter with addition
63 - 68  unused
### B.3. PIN ASSIGNMENTS FOR QED CHIP

<table>
<thead>
<tr>
<th>Pin</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2, 23, 24</td>
<td>$V_{SS}$</td>
</tr>
<tr>
<td>2, 4, 5, 10,</td>
<td>Unused</td>
</tr>
<tr>
<td>3, 6, 9, 14</td>
<td>Ground</td>
</tr>
<tr>
<td>7 - 8</td>
<td>Bias voltage to q17</td>
</tr>
<tr>
<td>11, 12, 33, 34</td>
<td>$V_{CC}$ with 5 V</td>
</tr>
<tr>
<td>13, 19</td>
<td>Bias voltage for digital tuning</td>
</tr>
<tr>
<td>15</td>
<td>Bias voltage for analog tuning</td>
</tr>
<tr>
<td>16, 18</td>
<td>Differential inputs</td>
</tr>
<tr>
<td>17, 20, 25, 28</td>
<td>Ground</td>
</tr>
<tr>
<td>21</td>
<td>Bias voltage to q7 and q8</td>
</tr>
<tr>
<td>24, 26, 27</td>
<td>$V_{CC}$ with 10 V</td>
</tr>
<tr>
<td>29, 30, 32, 35</td>
<td>Unused</td>
</tr>
<tr>
<td>31, 36, 39, 42</td>
<td>Ground</td>
</tr>
<tr>
<td>37, 41, 43</td>
<td>Unused</td>
</tr>
<tr>
<td>38, 40</td>
<td>Differential outputs</td>
</tr>
</tbody>
</table>