AFRL-RY-WP-TM-2009-1127

INTEGRATED MICRONODE RESEARCH

Kenneth K. O, Joe E. Brewer, Robert M. Fox, Rizwan Bashirullah, and John G. Harris

University of Florida

APRIL 2009
Final Report

Approved for public release; distribution unlimited.

See additional restrictions described on inside pages

STINFO COPY

AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7320
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the USAF 88th Air Base Wing (88 ABW) Public Affairs Office (PAO) and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RY-WP-TM-2009-1127 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

*//Signature//

TONY K. QUACH, Project Engineer
Advanced Sensor Components Branch
Aerospace Components Division

//Signature//

BRADLEY J. PAUL, Chief
Advanced Sensor Components Branch
Aerospace Components Division

//Signature//

TODD A. KASTLE, Chief
Aerospace Components Division
Sensors Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

*Disseminated copies will show "//Signature//" stamped or typed above the signature blocks.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS**.

1. REPORT DATE ***(DD-MM-YY)***
April 2009

2. REPORT TYPE
Final

3. DATES COVERED ***(From - To)***
04 June 2008 – 03April 2009

4. TITLE AND SUBTITLE
INTEGRATED MICRONODE RESEARCH

5a. CONTRACT NUMBER

5b. GRANT NUMBER
FA8650-08-1-7854

5c. PROGRAM ELEMENT NUMBER
63739E

6. AUTHOR(S)
Kenneth K. O, Joe E. Brewer, Robert M. Fox, Rizwan Bashirullah, and John G. Harris

5d. PROJECT NUMBER
ARPR

5e. TASK NUMBER
YD

5f. WORK UNIT NUMBER
ARPRYD0F

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Florida
Office of Engineering Research
339 Weil Hall
Gainesville, FL 32611

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Research Laboratory
Sensors Directorate
Wright-Patterson Air Force Base, OH 45433-7320
Air Force Materiel Command
United States Air Force

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL/RYDI

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RY-WP-TM-2009-1127

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

13. SUPPLEMENTARY NOTES
PAO Case Number: 88ABW 2009-2576; Clearance Date: 15 Jun 2009. This report contains color.

14. ABSTRACT
A brief focused research effort was performed to examine three critical aspects key to the feasibility of very small single chip CMOS communication nodes called “micronodes.” Simulation and analyses were carried out for node operation at 2.4 GHz or 5.2 GHz to determine the practicality of on-chip antennas, adequately stable on-chip frequency references, and sufficiently low power dissipation to permit node operation with small coin batteries. The design goal was to achieve 20 m node to node and 1 km node to base station operation for µNodes approximately the size of an M&M™ candy. On-chip monopole antenna configurations simulated at 5.2 GHz were shown to be feasible using simple geometries, while 2.4 GHz monopole antennas required more exotic configurations. Using a technique called differential chip detection; simulation showed that adequately stable frequency references could be established to support operation over the temperature and voltage operating ranges. Trial designs and simulations were performed to determine whether 5 mW operating power levels could be achieved. The results slightly exceeded the 5 mW level, but were close enough to confirm the technical feasibility of optimized designs reaching sub- 5 mW levels.

15. SUBJECT TERMS
Communication node, mote, on-chip antenna, differential chip detection, frequency reference, on-chip oscillator

16. SECURITY CLASSIFICATION OF:

a. REPORT	b. ABSTRACT	c. THIS PAGE
Unclassified	Unclassified	Unclassified

17. LIMITATION OF ABSTRACT:
SAR

18. NUMBER OF PAGES
64

19a. NAME OF RESPONSIBLE PERSON (Monitor)
Tony K. Quach

19b. TELEPHONE NUMBER *(Include Area Code)*
N/A

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

CONTENTS

Figures v

Tables vi

ACKNOWLEDGEMENT vii

DISCLAIMER vii

SUMMARY 1

Antenna Summary ..... м

Frequency Reference Summary ..... н

Power Minimization Summary ..... н

INTRODUCTION ..... о

Seedling Objectives/Focus ..... о

Micronode Research Vehicle ..... п

METHODS, ASSUMPTIONS, AND PROCEDURES ..... р

ANTENNAS ..... с

Off Chip Antennas ..... у

Commercial ceramic antennas ..... у

Antennas imbedded in a printed circuit board 1л

On-Chip Antennas Operating at 5.2GHz 1н

Antennas with a single layer dielectric coating 1н

On-Chip Antennas Operating at 2.4GHz ..... мп

Antennas with two layer dielectric coating ..... мп

Antennas with inductive loading ..... мт

Antennas with both capacitive and inductive loading ..... му

On-Chip Antennas Operating at 60 GHz ..... мф

ON-CHIP FREQUENCY REFERENCE 2о

Crystal Avoidance Challenge 2о

Differential Chip Detection Approach 2о

Baseline PHY overview ..... нп

Baseband receiver architecture ..... нр

Digital baseband processing and design ..... нс

Differential chip decoder and inner correlator implementation ..... нс

Transmitter architecture ..... нф

On-Chip Reference Oscillator Approach 3л

POWER MINIMIZATION 3м

Power Source Issues 3м

Block Shut-Down Strategy 3м

Power Dissipation Baseline Targets and Simulation Results 3н

Differential Versus Single Ended Architecture 3о

Receiver Design ..... оп

Transmitter Design ..... от

injection locked 4.8 GHz frequency divider ..... оу

Phase multiplexer and interpolator ..... оу

Class-D power amplifier ..... оф

Transmitter simulation results 4л

Frequency Reference Power 4л

Digital Power 4м

Analog Processing Alternatives ..... пп

Summary of analog alternative approaches ..... пп

Research performed ..... пр

Iterative decoders ..... пр

A decoder example ..... пт

Issues at 65nm for analog iterative decoders ..... пу

DCD and DSSS in analog ..... пф

Future work ..... пф

REFERENCES ..... пф

FIGURES

Figure 1: µNode marble assembly concept ..... ф. р

Figure 2: Simplified antenna equivalent circuit for transmit mode ..... ф... т

Figure 3: Chip antenna and recommended PCB. ..... ф.. у

Figure 4: Chip antennas with different PCB sizes and return loss characteristic. ..... ф... ф

Figure 5: Antenna pair gain (G_a) measurement setup and result. ..... ф... ф

Figure 6: Three printed circuit board antenna simulation models 1л

Figure 7: µNode marble elements 1н

Figure 8: On-chip antennas with single-layer dielectric coating. 1о

Figure 9: On-chip antennas with double-layer dielectric coating. ..... мр

Figure 10: On-chip antennas with spherical coating. ..... мс

Figure 11: Antennas with inductive loading. ..... мт

Figure 12: Slow-wave antenna. ..... му

Figure 13: Cross section of patch antenna formed on a CMOS chip ..... мф

Figure 14: Preliminary design of the on-chip patch antenna 2л

Figure 15: Final design of the on-chip patch antenna. 2л

Figure 16: HFSS simulation results: (a) Input return loss (S_{11}) (b) Smith chart (S_{11}) (c) Radiation patterns (d) 3D polar radiation plot. 2м

Figure 17: Back-end process flow for the on-chip patch antenna 2н

Figure 18: Differential chip detection block diagram ..... нп

Figure 19: Block diagram for M-ary quasi-orthogonal modulator. ..... нп

Figure 20: Block diagram for the µNode Receiver ..... нр

Figure 21: Block level diagram of the µNode digital signal processor ..... нс

Figure 22: Differential chip detection (a) functional schematic and (b) timing waveforms ..... нт

Figure 23: Correlator Peak detection ..... ну

Figure 24: Memory Architecture of Outer Correlator ..... нф

Figure 25: Transmitter Block 3л

Figure 26: Power system for µNode 3м

Figure 27: Differential RF front-end µNode 3о

Figure 28: Single-ended RF front-end µNode ..... оп

Figure 29: Receiver block diagram ..... ор

Figure 30: One half of the receiver front end circuit ..... ор

Figure 31: VGA basic cell ..... ос

Figure 32: Transmitter OQPSK (MSK) direct digital modulator ..... от

Figure 33: Injection locked frequency divider (a) and delay cell detail (b) ..... оу

Figure 34: Phase multiplexers (a) Type-I, (b) Type-II ..... оф

Figure 35: Phase Interpolator ..... оф

Figure 36: Class-D Power Amplifier 4л

Figure 37: µNode digital signal process 130 nm CMOS layout 4м

Figure 38: Digital power distribution 4н

Figure 39: Replacing digital decoder with analog decoder ..... пс

TABLES

Table 1: Free space wavelengths versus frequency ..... фф... т

Table 2: PCB antenna models and simulation results 1л

Table 3: On-chip antenna with single layer dielectric simulation results 1н

Table 4: On-chip antenna simulation results for antenna with double-layer dielectric coating ..... мр

Table 5: Antenna gain for high permittivity first coating and spherical second coating ..... мт

Table 6: Simulation results for antennas with inductive loading. ..... мт

Table 7: Simulation results at 2.4 GHz for slow-wave antennas..... му

Table 8: Patch antenna final design parameter and simulation results 2м

Table 9: Differential Chip Encoding/Decoding..... ну

Table 10: Summary of µNode operating power dissipation 3н

Table 11: Receiver simulation results ..... ос

Table 12: Transmitter simulation results 4л

Table 13: Power Consumption of each block 4л

Table 14: Digital power estimates for 130nm CMOS process 4н

Table 15: Power comparison for various digital logic blocks in 130nm and 65nm 4о

Table 16: Gate capacitance per unit area in 130nm and 65nm..... 4о

Table 17: Capacitance per unit micron of transistor width in 130nm and 65nm..... пп

Table 18: Short circuit current in 130nm and 65nm..... пп

Table 19: I_{OFF} and Gate leakage per unit area in 130nm and 65nm.....пп

Table 20: Fabricated, measured and reported digital iterative decoders ..... пс

Table 21: Fabricated, measured and reported analog iterative decoders ..... пт

ACKNOWLEDGEMENT

This material is based on research sponsored by Air Force Research Laboratory (AFRL) and the Defense Advanced Research Agency (DARPA) under agreement number FA8650-08-1-7854. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation thereon.

DISCLAIMER

The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of Air Force Research Laboratory (AFRL) and the Defense Advance Research Agency (DARPA) or the U. S. Government.

SUMMARY

A focused research effort was conducted to examine the technical feasibility of on-chip antennas, on-chip frequency references and the achievement of adequate power minimization. These three areas were identified as critical for implementation of large scale integrated RF systems on silicon.

ANTENNA SUMMARY

Technical feasibility of antennas of dimensions small compared to the wavelength of the operating frequency was investigated. The study targeted alternative antenna structures that could fit into a µNode marble geometry and operate at either 5.2 GHz or 2.4 GHz. Both off-chip and on-chip options were examined. The criteria for judging the adequacy of a particular antenna configuration was achievement of approximately -2 dBi of antenna gain and an input resistance of about 10 ohms or higher.

Off-chip commercial ceramic antennas were found to require mounting on a printed circuit board to be usable, and the dimensions of the antenna assembly were incompatible with µNode packaging goals. The antenna performance was found to be significantly below (~ 6 dB) that given in the manufacturers data sheet, and was very sensitive to the PCB configuration. Off-chip antennas embedded in a printed circuit board were more promising, and may be compatible with some forms of µNode packaging. At 5.2 GHz a simple monopole offered 1.1 dBi of gain and 7.35-j45 ohm input impedance – a usable performance level. At 2.4 GHz the gain was -4.6 dBi and the impedance was 2.5-j204 – both parameters lower than desired. An experiment using a helical shape for the antenna was successful in raising the impedance 22+j225, but it did not alter the gain. Like the ceramic antennas, antennas embedded in a PCB are expected to be sensitive to the board configuration.

On-chip antennas were investigated by constructing 3D models of proposed structures and evaluating gain and input impedance at the two frequencies of interest. Antenna variations included simple metal on the top surface of a chip, and metal with an etched out air gap under the antenna. The impact of encapsulating materials were investigated for the cases of a single layer coating with a relative dielectric constant of 4, and a double layer coating where the first layer had a dielectric constant of 10 and the second layer had a constant of 4.

At 5.2 GHz a 7 mm long monopole with a single layer coating was observed to exhibit a gain of -3.8 dBi and impedance of 22.3-j15 ohms. When a 100 thick air gap was introduced underneath the antenna, the gain increased to +0.2 dBi and the impedance became 7.9-j96. This confirmed the feasibility of on-chip antennas for the 5.2 GHz frequency.

At 2.4 GHz the situation proved to be more complex. Single layer coatings over a 7 mm monopole without an air gap provided -17.9 dBi and 40.7-j158. Introducing a 100 micron air gap resulted in -9.6 dBi and 8.4-j330. Two layer coatings were tried, and caused the gain to increase and the impedance to reduce. One two layer simulation yielded -6.0 dBi and 2.2-j113, while a second one yielded -3.5 dBi and 2.8-j94. None of these configurations met the -2 dBi and 10 ohm goal.

The 2.4 GHz investigation was expanded to include the options of loading the antenna in various ways. An inductive loading approach yielded -7.7 dBi and 4.8-j260 ohms. Combined inductive and capacitance resulted in -18.26 dBi and 170-j194 ohms. It was concluded that achievement of the desired gain and input resistance within the size constraint would require considerable optimization that must combine both material and geometric variables. This work was beyond the scope of this seedling effort.

The option of using on-chip patch antennas at 60 GHz was briefly examined. The study treated a simple patch structure that employed BCB as the dielectric between the patch and the ground plane. An exploratory simulation study investigate a particular geometry with varied dielectric thickness, and noted that performance improved as the thickness increased. Using that initial simulation information a specific patch geometry was defined and simulated in some detail. The resultant gain was 4.9 dBi with an input resistance of 178 ohms. The research provided illustrative diagrams of the patch structure on the top surface of a chip, and showed the nature of the post CMOS processing necessary to form the patch. It was concluded that on-chip patch antennas operating at high frequencies were not only technically feasible, but are capable of superior performance.

FREQUENCY REFERENCE SUMMARY

Node to node communication requires that the transmitting frequency and the receiver capability of decoding the transmission be compatible. In general, the frequency references on each chip must be close to the same frequency and stable over the operating temperature and supply voltage variation ranges. The usual way of coping with this restriction is to use an off-chip crystal to stabilize each oscillator. In the µNode the size and cost of an off-chip crystal cannot be tolerated. A technique called "differential chip detection" (DCD) was used in a 24 GHz µNode design to relax the required stability specification for the reference oscillator so that the option of using an on-chip oscillator could be considered. This research examined the feasibility of DCD at 5.2 and 2.4 GHz by designing and simulating a digital processor to perform this function. The design relaxed the frequency requirement to the point that ±100 ppm frequency variation could be allowed.

An on-chip reference oscillator design was performed and simulated using 130 nm CMOS. The design featured resources that facilitated a onetime calibration of the frequency within ±10 ppm, and provided a dynamic correction of variations due to temperature such that ±100 ppm could be readily achieved. The temperature compensation scheme would also support achievement of ±25 ppm stability if the stored compensation curve were more elaborate. A final simulation of the overall reference system confirmed the capability of the system to achieve the required stability, and it was concluded that the on-chip reference was technically feasible.

POWER MINIMIZATION SUMMARY

Studies were performed to determine the feasibility of a 5.2 GHz or 2.4 GHz µNode design that could operate in receive mode or transmit mode with a power dissipation of 5 mW when implemented using

65 nm CMOS. Trial designs were prepared for all of the functional blocks within a µNode and simulations were performed to obtain power estimates. For some parts of the µNode the procedure was complicated by the lack design libraries for the 65 nm process, and 130 nm simulations had to be substituted. It was then required that the 130 nm estimates be scaled to represent what could be achieved at 65 nm. The sum of all the individual block power estimates was 5.2 mW for receive and 5.0 mW for transmit. Since the circuit designs prepared for this study were not optimized, it was concluded that achievement of 5 mW in a future design was technically feasible.

INTRODUCTION

The Integrated Micronode (µNode) program was awarded 4 June 2008 to the University of Florida (UFL) Silicon and Microwave Integrated Circuits and Systems (SiMICS) group in the Department of Electrical and Computer Engineering for the purpose of rapidly exploring key feasibility issues associated with establishment of self contained RF subsystems in silicon technology.

SEEDLING OBJECTIVES/FOCUS

Three critical feasibility issues were to be examined: antenna implementation, communication frequency reference implementation, and power minimization.

Antennas are a necessary part of all RF communication subsystems, and antenna realization in such a manner as to avoid wired connections that go off-chip would be a definite plus. In addition, practical on-chip antennas would facilitate realization of systems that require multiple antennas such as ULSI imaging devices. This research examined the capabilities of on-chip antennas formed using normal metallization layers available in standard CMOS processes in combination with inexpensive post CMOS fabrication steps such as addition of thicker metal and dielectric layers, and silicon etching. Another aspect of the antenna research was to examine the nature and limitations of small off-chip antennas, and compare that option with the capabilities of on-chip antennas.

When RF subsystems communicate with each other they typically share a common carrier frequency, and it is necessary that the frequencies generated on the separate radios be very close to one another. A common solution is the utilization of off-chip crystals to force the separate oscillators to establish and maintain frequency references within a tight tolerance. For subsystems with a minimum cost and physical volume objective, the use of such off-chip components is objectionable. For systems that may employ a large number of communication nodes, the cost difference may make many applications unaffordable. For systems that are intended to be semi-covert, the size difference may be such that the node becomes obvious to a casual observer. It was one objective of the research to explore the potential of a technique called “differential chip detection” to relax frequency tolerance requirements to such an extent that on-chip, non-crystal stabilized, oscillators can be used to implement an adequate reference.

Power minimization is a desirable universal goal for all battery powered devices. The classic trade between unit size, available operating time, and choice of battery, dominates designs where small size is

desirable. It was an objective of the research to explore power minimization within the essential functional blocks of a practical silicon RF transceiver.

The overall goal of the research was to provide a factual basis for judging the technical and practical feasibility of realizing complete RF subsystems (or multiple subsystems) on single silicon chips.

MICRONODE RESEARCH VEHICLE

The seedling objectives were to explore the feasibility of designing and fabricating practical antennas, frequency reference circuits and very low power integrated RF subsystems on silicon. Feasibility implied more than working thorough just one special case design, so it was important that the research vehicle allowed verification of capabilities that are applicable to a broad range of RF integrated circuits on silicon. A single chip ultra small communication node was believed to have such general validity. The device was called a micronode (µNode).

A µNode requires all elements of RF transmitter and receiver subsystems including frequency reference circuits and digital baseband signal processing circuits. Because it is implemented on mainstream affordable CMOS it offers low cost and high levels of integration. This particular circuit has the additional advantage of offering a level of performance that can be exploited in a number of vital military applications including sensor networks and short range communication aids.

In this seedling effort the criterion for achievement of “feasibility” was the demonstration through simulation and analyses that µNode CMOS designs can provide acceptable performance in the end product. “Acceptable performance” is, of course, not an absolute thing. It can only be defined in the context of the application requirements. Specifically, UFL proposed to use single chip micronodes (µNodes) operating at 2.4 GHz or 5.2 GHz as the vehicle for the research. A µNode device, the approximate size of an M&M™ candy, capable of node to node communication distances of 20 m, and node to base station communication distances of 1 km was the target.

Work at UFL over the past decade established a significant technical base for µNode technology. All of the functional blocks required for a µNode operating at 24 GHz were previously defined and demonstrated in 130 nm CMOS. For the purposes of the research reported here, major portions of the existing 24 GHz design were scaled down to 2.4 GHz and retargeted from 130 nm CMOS to 65 nm CMOS. The 10x frequency reduction allowed the use of much simpler radio architecture with a resultant power and chip area saving.

The target size of the µNode assembly was a sphere with approximately 11.5 mm diameter, 0.8 cm^3 volume and 0.8 grams mass which was just about the size of an M&M™. The target power dissipation was 5 mW when operating. More than one year of life was projected when operated at a 0.1% duty cycle. The desired procurement price was less than $2. A concept sketch for one packaged version of a µNode assembly was presented in Figure 1. Here the µNode chip with on-chip antenna was mounted on top of a battery. The off-center mass of the battery within the sphere causes the marble to self orient with the antenna pointing upward.

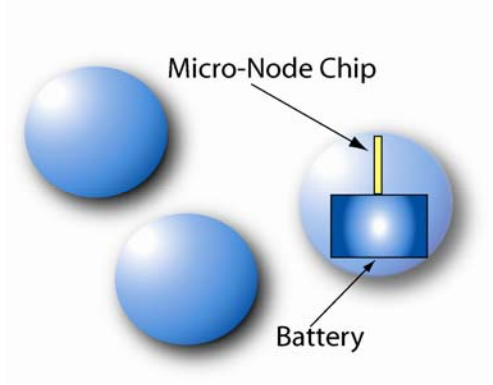


Figure 1: µNode marble assembly concept

METHODS, ASSUMPTIONS, AND PROCEDURES

The research goal was to examine the technical feasibility of only three aspects of µNode technology in a rapid response manner with a minimum expenditure of funds. The investigations were limited to simulations and analyses. No integrated circuit fabrication or extensive experimental activities were undertaken. Samples available from previous research were made use of where available, and some inexpensive measurements were performed where appropriate. The primary antenna modeling tool was HSSS from Ansoft. The 3D full-wave electromagnetic field simulation HFSS™ tool provided by the Ansoft division of Ansys, Inc. was used to simulate alternative antenna configurations.

ANTENNAS

The function of an antenna is to radiate or receive electromagnetic waves. The µNode goal of achieving dimensions approximately the size of and M&M™ candy forces the antenna to be small, and such an antenna necessarily has a small capture cross section. While an antenna can be formed by metal structures either on or off the silicon chip, the lowest cost and smallest overall assembly size is achieved by using on-chip structures formed using the normal metallization layers available in a mainstream CMOS process. On the other hand, an on-chip antenna must operate in the presence of losses associated with the silicon substrate and thin metal lines; a problem that off-chip antennas can to some degree avoid. The technical feasibility of an on-chip antenna operating at 5.2 GHz or 2.4 GHz depends on achievement of adequate antenna gain and reasonable impedance match between the antenna and transceiver circuits. Antenna gain can be expressed as: $G = 4 \pi \eta A/\lambda^2$. Here G is the antenna gain, η is antenna efficiency, A is the effective antenna area that is dependent on direction and λ is the wavelength of the signal.

For the µNode application, system level studies established -2 dBi as a criterion for acceptable on-chip antenna gain. That is, the antenna was required to exhibit gain no worse than 2 dB below that of a lossless isotropic antenna. Achievement of adequate gain requires a good level of understanding of the factors contributing to antenna losses and the impact of antenna structural features.

The requirement for antenna input impedance was somewhat flexible in that it was associated with the capabilities of the transmitter power amplifier. A real value for the antenna input impedance of about 10 ohms or higher was desirable in order to achieve a reasonable impedance match. The simulation tool employed in this investigation automatically provided scattering parameter S_{11} and input impedance. Figure 2 presents a simple equivalent circuit for the antenna where conduction and dielectric losses have been lumped together as a single resistor R_{cd}, and the equations for input impedance and $|S_{11}|$ are summarized. The present assumption is that the transmitter source impedance is 50 ohms, but that impedance will actually be determined later when the chip design is finalized.

One important aspect of the feasibility of on-chip antennas was the relative size of the antenna versus the signal wavelength. Obviously it was desirable that the electrical length of antenna be comparable to a half or quarter wavelength so that the natural resonances of the antenna could be exploited. Free space wavelength is simply $\lambda_0 = c/f$ where c is the speed of light and f is the frequency. Table 1 shows that at the 2.4 and 5.2 GHz frequencies the quarter wavelength approaches practical dimensions for on-chip antennas. At higher frequencies, such as 60 or 90 GHz, quarter wave antennas are certainly compatible with practical chip dimensions. Of course, the free space wavelengths can be adjusted by inserting materials with a high dielectric constant around the antenna (i.e. $\lambda \approx \lambda_0/\sqrt{\varepsilon_r}$), and this was one approach that was considered. In any case, in the normal process of packaging a µNode some materials with a dielectric constant greater than air (typically about 4) would be used to protect the assembly, and even higher permittivity materials can be added to further adjust the effective wavelength.

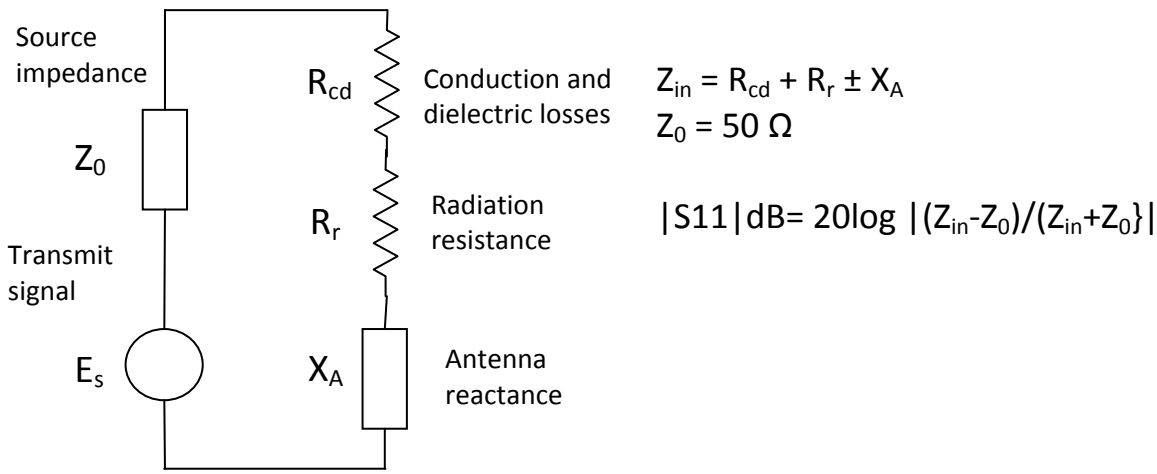


Figure 2: Simplified antenna equivalent circuit for transmit mode

Table 1: Free space wavelengths versus frequency

Frequency (GHz)	Wavelength λ_o (mm)	Quarter Wave λ_0 (mm)
1.8	166.7	41.67
2.4	125	31.25
5.2	57.7	14.425
10.0	30	7.5
60.0	5	1.25
90.0	3.3	0.833

In this research several approaches were studied to reduce antenna size while maintaining critical performance. As a starting point, some external (off-chip) commercial chip antennas were studied and characterized. Specifically, ceramic antennas that can be attached to a substrate were examined. Another off-chip antenna approach, formation of the antenna within a printed circuit board using normal metal line and via technology, was also investigated. The research then proceeded to use simulation to investigate alternative ways of preparing on-chip antennas.

The issue that dominated this antenna study was how to get acceptable performance from an antenna that could fit into the M&M™ size target for a µNode. Several common approaches to reducing antenna size were studied. One method was to introduce dielectric loading to adjust the wavelength -- both single and double dielectric layers. First the single-layer dielectric coating for on-chip antennas was investigated and found to be satisfactory for operation at 5.2 GHz, and then the more difficult problem of working at lower frequencies was treated. Operation at 2.4 GHz was examined using double-layer dielectric coatings. In addition, other techniques, such as schemes for loading the antenna inductively or loading using both inductive and capacitive elements (i.e. slow wave structures), were investigated. Finally the research briefly examined some aspects of antennas operating at higher frequencies such as 60 GHz.

OFF CHIP ANTENNAS

COMMERCIAL CERAMIC ANTENNAS

Off the shelf commercial antennas are available for operation at 2.4 GHz, but only a few of them approach the small size needed for the µNode application. Even when the antenna itself is quite small, the device must be attached to a printed circuit board (PCB), and the combined assembly size is in conflict with the size objective of the µNode.

The chip antenna, AN3216, from the Rain Sun Company was procured, studied and characterized. Figure 3 shows the chip antenna and a manufacturer recommended PCB. While the chip antenna had a reasonable size, the 35 x 50 mm PCB was objectionable.

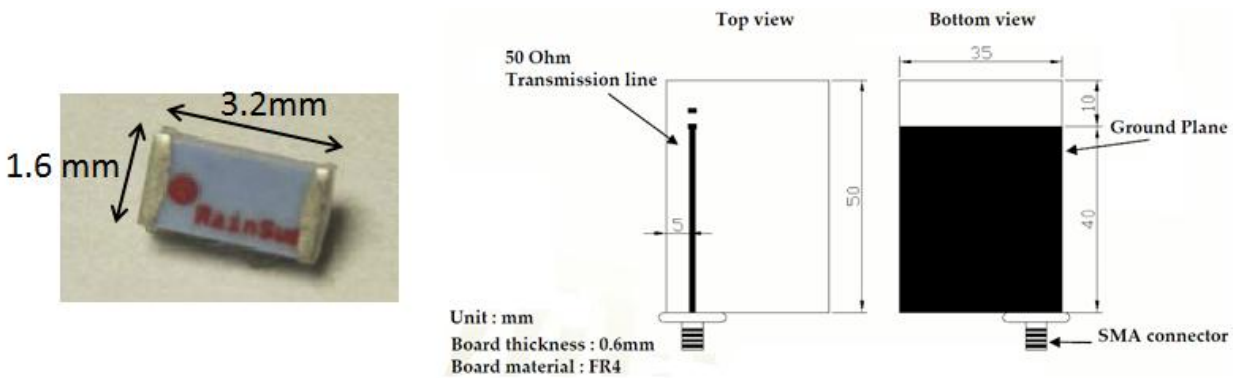


Figure 3: Chip antenna and recommended PCB.

The electromagnetic influence of PCB size was investigated by preparing several alternative boards and performing measurements and simulations. Figure 4 shows antennas with different PCB sizes and the return loss characteristic of each one. As indicated in the figure, both tuning frequency and return loss vary with the overall size of the PCB. Because a large PCB was required to achieve the desired operating characteristics at 2.4 GHz this approach was not suitable for use as part of a small µNode.

Additional measurements were performed to explore the antenna-PCB assembly performance. Figure 5 shows the antenna pair gain (G_a) for these devices. Also, the antenna gain was estimated from this measurement result. In this figure, two antenna-PCB samples (sample #1 and #4) were used for comparison.

The antenna pair gain (G_a) was computed using equation 1-1.

$$G_a = \frac{|S_{21}|^2}{(1-|S_{11}|^2)(1-|S_{22}|^2)} = G_t G_r \left(\frac{\lambda}{4\pi R}\right)^2 \qquad (1\text{-}1)$$

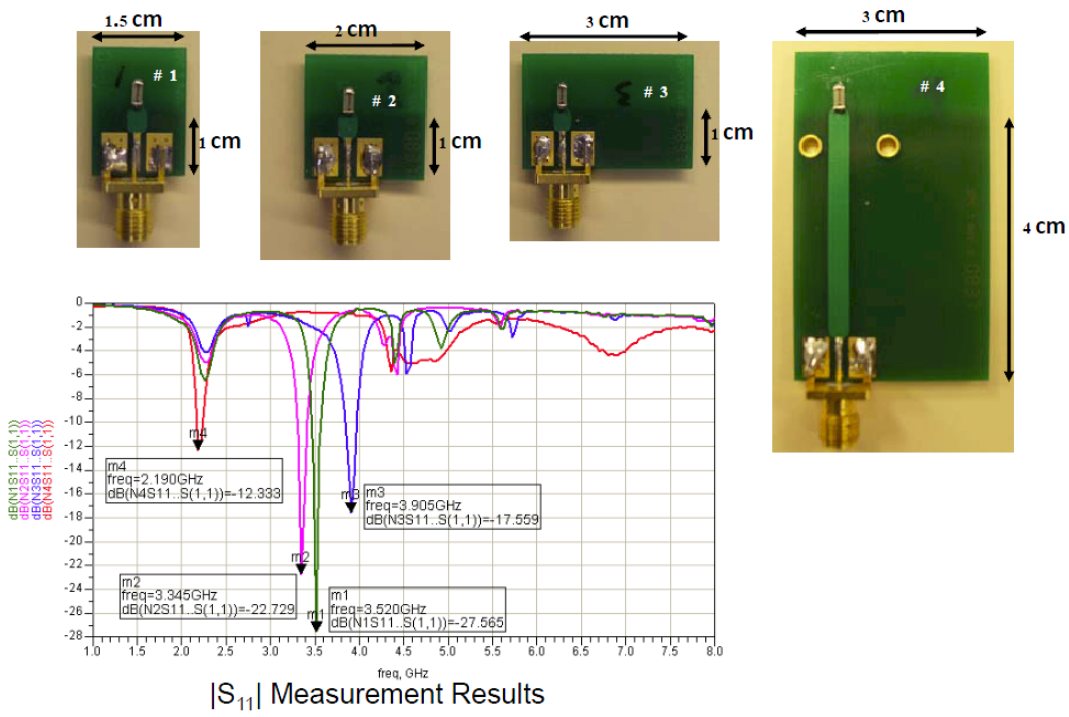
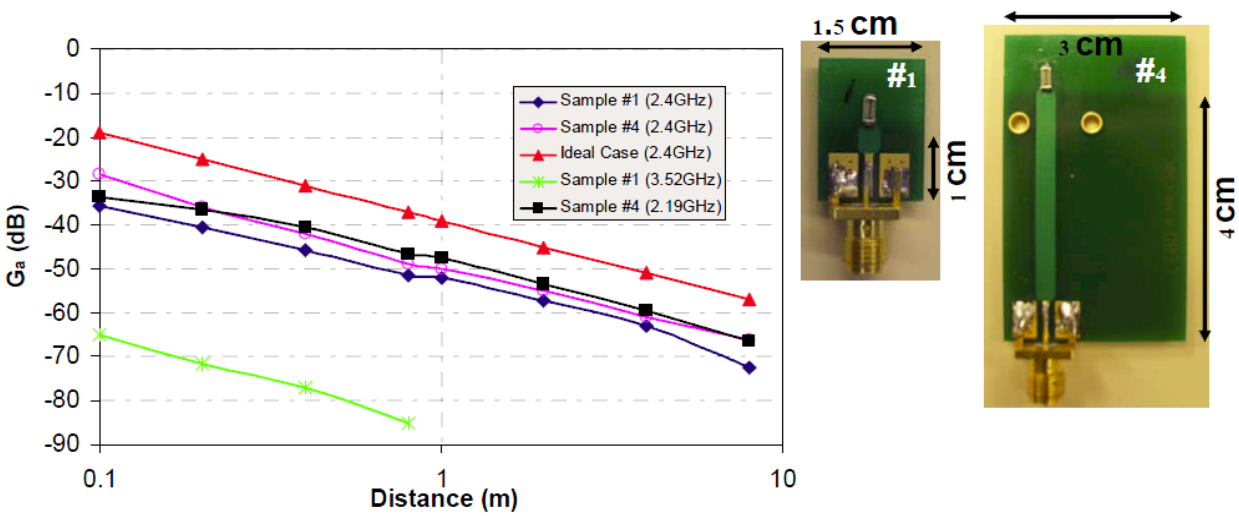


Figure 4: Chip antennas with different PCB sizes and return loss characteristic.



	Gain (dBi) @ 2.4 GHz	Gain (dBi) @ Tuned Frequency
#1	- 6.8	-20.8 (@ 3.52GHz)
#4	-5.5	-4.6 (@ 2.19GHz)

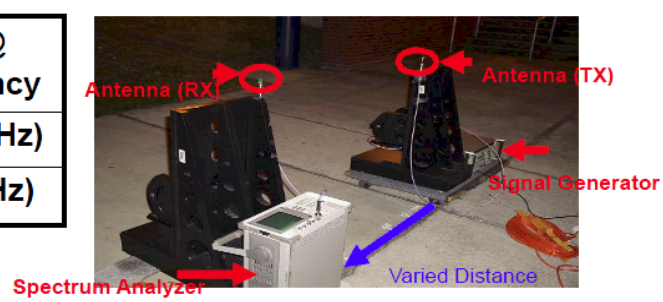


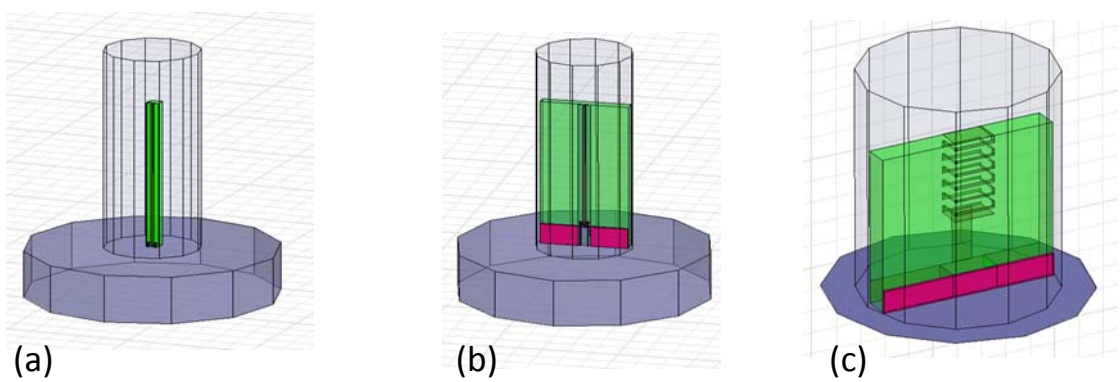
Figure 5: Antenna pair gain (G_a) measurement setup and result.

The measured loss due to the fixtures was added back to the overall gain measurements to correct for set-up loss. It was observed that the antenna pair gain (G_a) of the antenna was lower than the ideal case specified in the datasheet. This further confirmed the conclusion that this class of antenna would not be suitable for use in small µNode devices.

ANTENNAS IMBEDDED IN A PRINTED CIRCUIT BOARD

Metal line and via technology can be used to form an antenna structure on a printed circuit board. The merit of using a PCB antenna in a µNode application depends to some extent on the nature of the particular application. If a particular sensor or other component is required that cannot be accommodated on the µNode chip, a PCB of some type may be needed to serve as the base of the assembly. In this case, the PCB is an essential overhead item, and using it as the antenna platform comes almost for free. The PCB antenna, of course, would not have silicon substrate losses as in the case of the on-chip antenna.

To examine the potential for the PCB alternative in the µNode marble configuration three simulation models were constructed as shown in Figure 6. Table 2 presents the structural details for each model and the simulation results at 5.2 GHz and 2.4 GHz.



(a) Model of on-chip monopole where silicon was replaced by FR-4

(b) PCB monopole model with chip mounted on the board

(c) PCB coil antenna model with chip mounted on the board

Figure 6: Three printed circuit board antenna simulation models

Table 2: PCB antenna models and simulation results

#	Antenna Metal			PCB			Encapsulation		5.2 GHz		2.4 GHz	
	Thick mm	Length mm	Width mm	Thick mm	Length mm	Width mm	Diam mm	Height mm	Gain dBi	Impedance ohms	Gain dBi	Impedance ohms
a	0.035	7	0.12	0.443	7.28	0.5	3.8	10	1.6	6.6-j32.5	-3.5	1.7-j171
b	0.035	6	0.12	0.443	7.28	3.5	3.8	10	1.1	7.35-j45	-4.6	2.5-j204
c	0.035	31	0.1	1	7	7	7.5	10	-3.7	4.9-j91.3	-4.8	22+j225

The model concept was to approximate the µNode electromagnetic environment by positioning the PCB perpendicularly to a metal cylinder that represented a coin battery. The PCB substrate was specified to

be FR-4, and its relative dielectric constant was taken to be 4. In addition to the PCB structure, the models incorporated a cylindrical solid that surrounded the PCB to approximately represent an encapsulating material.

Model (a) consisted of a simple 7 mm long straight wire monopole on the PCB surface encapsulated within a cylinder (3.8 mm diameter, 10 mm high) with an ε_r of 4. Since the encapsulating material was assumed to have the same properties as the PCB, the dimensions of the PCB are not relevant to the model (i.e. the encapsulation simply acts as an extension of the PCB material). At 5.2 GHz both antenna gain (1.6 dBi) and the real portion of the impedance (6.6 ohms) of this non-optimized structure appear to be usable. At 2.4 GHz the gain (-3.5 dBi) and the real portion of the impedance (1.7 ohms) do not look promising, and further optimization is needed.

Model (b) was intended to provide a more realistic representation of a µNode configuration. Here a silicon chip was included at the base of the antenna to represent the µNode device, and the width of the PCB was increased to 3.5 mm to provide a platform for the chip and a ground reference pattern. The length of the antenna metal was reduced to allow space for the die and metal and still remain within the µNode marble envelope. As a result of these changes the antenna gain dropped slightly (down 0.5 dB to 1.1 dBi at 5.2 GHz, and down 1.1 dB to -4.5 dBi at 2.4 GHz), and the real portion of the impedance increased slightly (up by 0.8 ohms to 7.4 ohms at 5.2 GHz, and up by 0.8 ohms to 2.5 ohms). As in the case of model (a), operation of this simple monopole at 5.2 GHz looks like it has some potential, but both the gain and the impedance at 2.4 GHz need a good bit of improvement. The small size of these antennas relative to the quarter wavelength prevents them from taking advantage of resonance effects.

Model (c) departed considerably from the simple straight wire approaches of the first two models. A helical shaped structure was used to increase the effective length of the antenna without requiring a longer PCB. The helical coil was formed by metal lines on both sides of the PCB connected by through the board via metal. The simulated structure was intended to approach resonance at 2.4 GHz which would occur in a folded structure at approximately a 31 mm total line length (approximately twice the 15.6 mm wavelength in an ε_r=4 material). The helical design did not change the 2.4 GHz gain very much (down by 0.2 dB from the (b) model to -4.8 dBi), but it did greatly improve the real portion of the input impedance (up by 19.5 ohms from the (b) model to 22 ohms).

While none of the models achieved the µNode design target of > -2 dBi gain combined with a real portion of the input impedance of about 10 ohms, all of them indicated that satisfactory operation could be achieved at 5.2 GHz with little optimization effort. The helical PCB antenna offers the best approach for achieving satisfactory operation at 2.4 GHz. The simulation confirmed that adequate input impedance can be achieved, and it is expected that some refinement in the definition of the helical coil can raise the gain.

ON-CHIP ANTENNAS OPERATING AT 5.2GHZ

The characteristics of on-chip antennas were explored for a specific geometric configuration and a specific antenna length. That is, the µNode marble configuration with the approximate size of an M&M™ candy was taken as defining the desired geometry. Figure 7 points out the critical elements within the marble that must be represented. A series of simulation models were constructed assuming a fixed length of 7 mm for the metal antenna, and an antenna location perpendicular to a metal cylindrical surface that represents the battery. In addition, the models included treatment of the impact of possible encapsulation materials that would fill the internal portions of the shell. The model structure included a battery (230-µm thickness, 1-cm diameter), a 7-mm linear monopole on-chip antenna with 30-µm metal width, and a 20-Ω-cm silicon substrate (100-µm thickness). The metal (Aluminum) thickness was 3 µm. The silicon dioxide thickness was 3 µm.

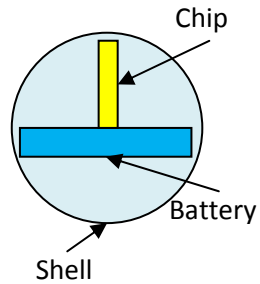


Figure 7: µNode marble elements

ANTENNAS WITH A SINGLE LAYER DIELECTRIC COATING

From the viewpoint of having a desirable antenna pattern, the simple monopole has many advantages for µNode type applications. Where ever the antenna can be oriented vertically relative to the ground plane, the monopole offers good range and a circular non-directional characteristic. Since the simplest monopole is a straight wire, it also is most economical in use of chip surface area. Over the past decade the UFL research team established a significant body of experimental and analytical data that confirmed excellent performance for on-chip monopoles in µNode type applications. To explore the range of performance achievable by a simple monopole within the dimensional constraints of a µNode marble geometry, six models were prepared and simulated at 5.2 and 2.4 GHz. Table 3 summarized the attributes of each model and presented the gain and input impedance data for each frequency. Figure 8 shows the structural details common to all the models.

Table 3: On-chip antenna with single layer dielectric simulation results

#	Metal	Antenna	Air Gap µm	Silicon Thickness µm	Encapsulation ε_r=4 H=8 mm	Gain dBi		Impedance ohms	
						2.4 GHz	5.2 GHz	2.4 GHz	5.2 GHz
1	PEC	7 mm 30 µm 3 µm	None	None	None	+1.3	+1.9	0.8-j811	4.9-j320
2	Al					-1	+1.1	1.4-j800	6-j320
3	PEC				D1=5 mm	-1.6	+1.7	0.9-j228	5.4-j32
4	Al			100 (20 Ω-cm)	D1=3 mm	-17.9	-3.8	40.7-j158	22.3-j15
5	Al		100		D1=2 mm	-9.6	+0.2	8.4-j330	7.9-j96
6	Al				D1=6 mm	-8.1	-0.1	4.3-j223	6.9-j55

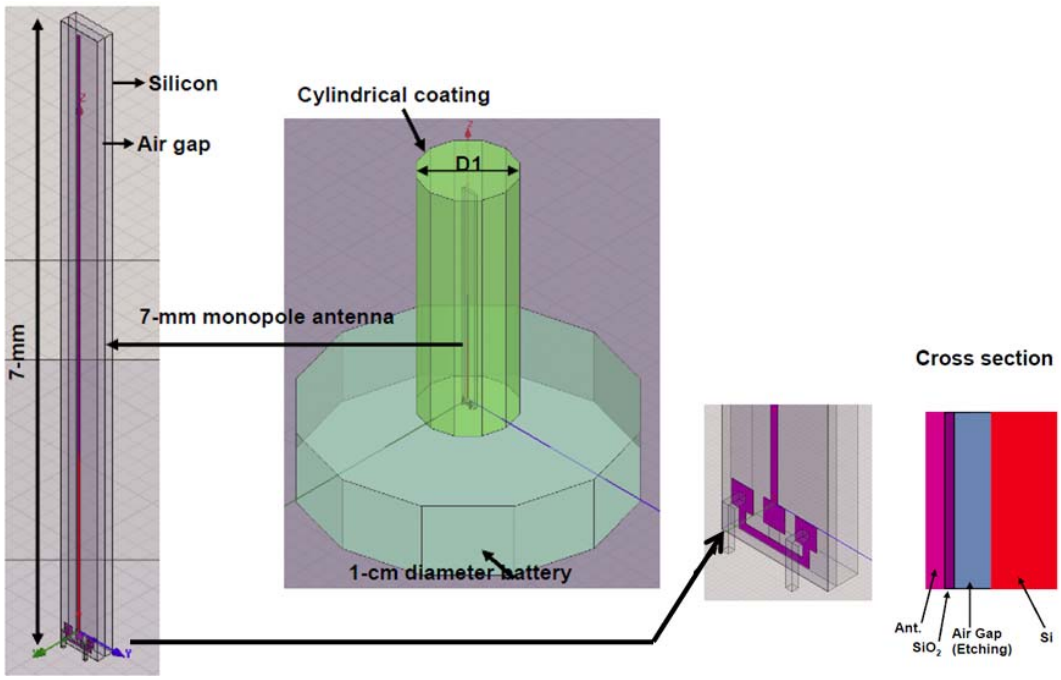


Figure 8: On-chip antennas with single-layer dielectric coating.

Model #1 was intended to establish an idealized baseline for the small antenna without any losses. The metal trace was specified to have zero resistivity (i.e. perfect electrical conducting (PEC) metal), and was hanging unsupported in free space (i.e. no air gap under the metal and no silicon substrate). The idealized antenna shows adequate gains but low input impedance at both frequencies

Model #2 introduced a realistic metal (aluminum) into the model configuration. As a result the gain degraded and the real portion of the impedance increased. The introduction of metal losses reduced the gain at 5.2 GHz by 0.8 dB down to +1.1 dBi while the real portion of the impedance increased by 1.1 ohms to 6 ohms.

Model #3 introduced an encapsulating material into the ideal Model #1 configuration. This 5 mm diameter cylinder of FR-4 (relative dielectric constant assumed to be 4) was expected to alter the effective wavelength thus changing any resonance effects, and to introduce losses in the material. Referring back to Table 1, the free space quarter wavelength at 5.2 GHz is about 14.4 mm. In a material of ε_r=4 the effective quarter wavelength should approach 7.2 mm. The simulation results at 5.2 GHz were +1.7 dBi and 5.4-j32 ohms. The encapsulation material losses degraded the gain by 0.2 dB and increased the real portion of the impedance by 0.5 ohms.

Model #4 was the first realistic configuration. Here aluminum metal, a silicon substrate and a coating material was included. The diameter of the coating cylinder was reduced to 3 mm from the 5 mm used in Model #3, so the results are not directly comparable. At 5.2 GHz the results were -3.8 dBi and 22.3-j15. This was lower gain than desired, but still a usable value that could readily be improved. Model #5

introduced an etched air gap underneath the antenna in order to reduce silicon losses. It also used a 2 mm diameter coating. The air gap improved the gain to +0.2 dBi and altered the impedance to 7.9-j96. This combination meets the µNode target of > -2 dBi and real portion of the input impedance of about 10 ohms. Model #6 repeated the conditions of Model #5 with an increase of the coating cylinder diameter from 2 mm to 6 mm. This change reduced the gain by 0.3 dB to -0.1 dBi, and altered the input impedance to 6.9-j55.

These simulation exercises for fixed antenna length with single layer coatings demonstrated the complex interplay between antenna gain and input impedance when measures are taken to reduce losses. Etching out the silicon substrate under the antenna improves gain but degrades the input impedance for matching purposes. The simulation results did confirm that a small simple monopole on-chip antenna can be made to work at 5.2 GHz in a µNode marble type configuration.

ON-CHIP ANTENNAS OPERATING AT 2.4GHZ

The six models of single layer encapsulated 7 mm long antennas were also simulated at 2.4 GHz, and as shown in Table 3 the antenna gains for the practical configurations (Models 4, 5 and 6) were far below the -2 dBi design target. It was concluded that measures beyond the single coating with an air gap would be required to achieve desired performance. The first option explored was the use of two layers of encapsulation for the purpose of altering the effective length of the antenna.

ANTENNAS WITH TWO LAYER DIELECTRIC COATING

A series of four models were constructed to examine the potential for double layer coatings as a method to achieve acceptable operation at 2.4 GHz. Figure 9 shows the structural details of the models. The concept was to use a first coating that has a high dielectric constant combined with a second coating that represented typical packaging encapsulants. The second layer also acts as an impedance transformer designed to maximize transmission and minimize reflections at the interfaces.

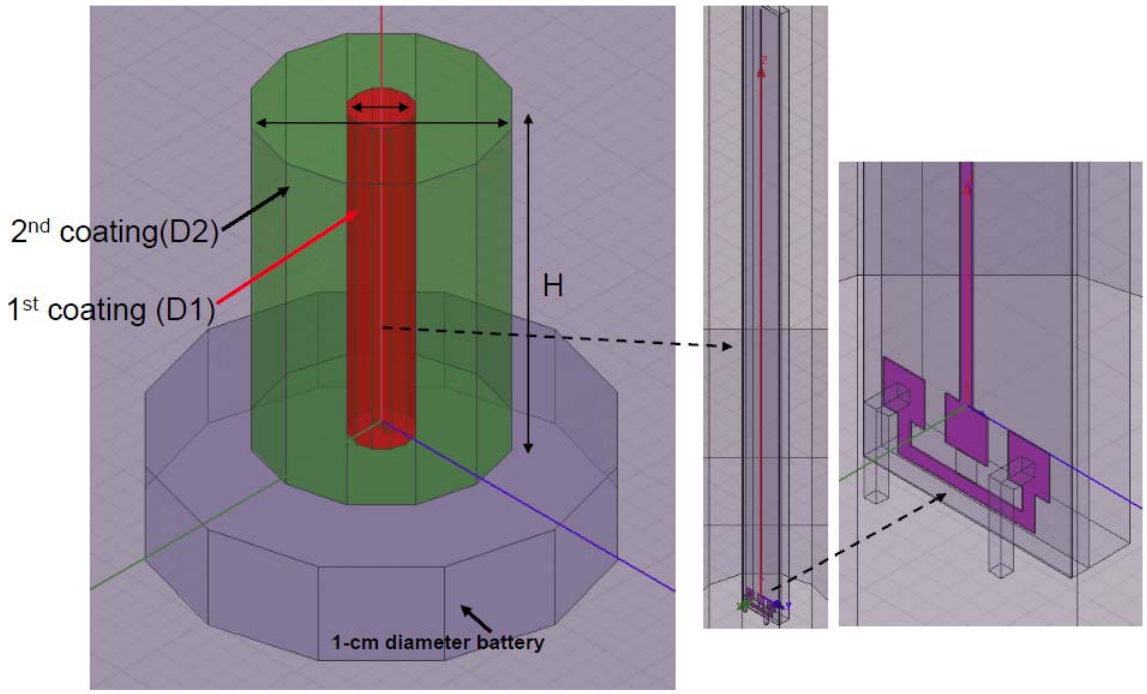


Figure 9: On-chip antennas with double-layer dielectric coating.

The structure includes a battery (230-µm thickness, 1-cm diameter) and a small linear type on-chip antenna with 30-µm metal width and 20-Ω-cm silicon substrate (100-µm thickness). The metal (copper) thickness is 3 µm. The silicon dioxide thickness is 3 µm. The first coating cylinder is the material with permittivity of 10 and diameter of 2.4 mm. The permittivity for second coating cylinder is 4 and the diameter is 7.8 mm. There is a 100-µm air gap underneath the silicon dioxide layer. Table 4 presents the simulation results.

Table 4: On-chip antenna simulation results for antenna with double-layer dielectric coating

#	Metal	Antenna mm	Air Gap µm	Si (20 Ω-cm) Thickness µm	Coating 1 ε_r = 10 mm	Coating 2 ε_r = 4 mm	Gain dBi	Impedance ohms
1	PEC	7	None	None	D1=2.4	D2=7.8 H=10	0.1	0.5-j97
2	Cu		100	100			-6.0	2.2-j133
3	PEC	10	None	None		D2=7.8 H=11.5	0.3	1.1-j67
4	C u		100	100			-3.5	2.8-j94

Model #1 showed the gain and impedance achievable for a 7 mm long antenna without metal and silicon substrate losses. The gain of 0.1 dBi was acceptable, but the 0.5-j97 ohm impedance was difficult to match. As shown in Model #2, including metal and silicon losses deteriorated the gain down to -6 dBi and increased the impedance to 2.2-j133. Note that the double coating was effective in raising the gain versus the gain achievable with one coating (i.e. -6 dBi double coating versus -8.1 and -9.6 for single coatings).

Models #3 and #4 examined the impact of increasing the antenna length to 10 mm. The height of the coatings had to be increased to 11.5 mm to accommodate the larger antenna. The idealized Model #3 showed a 0.3 dBi gain, and 1.1-j67 ohm impedance. Model #4 which included metal and silicon losses showed a gain of -3.5 dBi and impedance of 2.8-j94 ohms.

These four models confirmed that performance at 2.4 GHz could be improved by using a double coating approach, but the desired target gain of >-2 dBi was not achieved for the permittivity levels investigated.

Before concluding the investigation of double layer coating potential it was decided to examine the option of using a higher permittivity first coating and modification of the second level coating to represent the spherical shape of the µNode marble. Figure 10 shows the structure and Table 5 summarized the data.

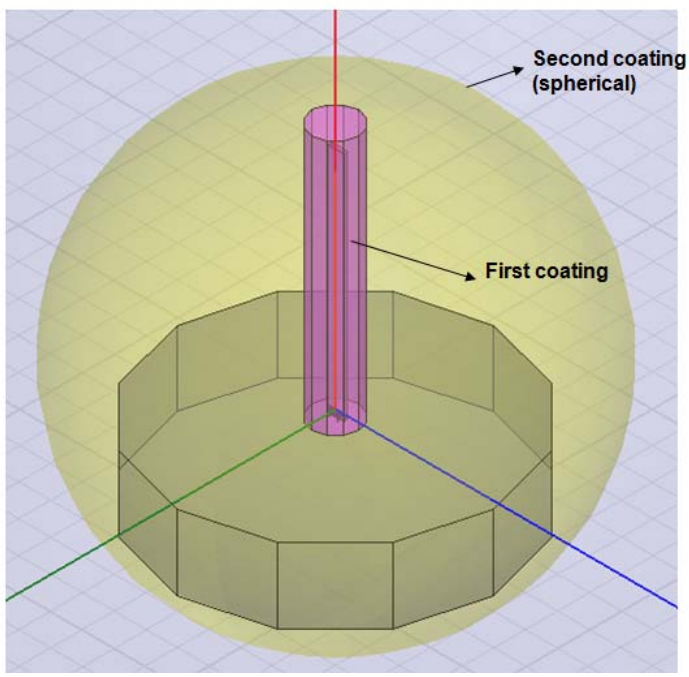


Figure 10: On-chip antennas with spherical coating.

This brief exercise showed a gain of -7.1 dBi at 2.4 GHz. Since this gain was slightly lower than the simpler lower dielectric ε_r (10 versus 50) of Model #2, this line of investigation was terminated in favor of more promising approaches.

Table 5: Antenna gain for high permittivity first coating and spherical second coating

#	Metal	Antenna mm	Air Gap µm	Si (20 Ω-cm) Thickness µm	Coating 1 cylinder ε_r = 50 mm	Coating 2 sphere ε_r = 2 mm	Gain dBi	Impedance ohms
1	Cu	7	40	100	D1=2.4	D=14	-7.1	2.2-j133

ANTENNAS WITH INDUCTIVE LOADING

Figure 11 shows two antenna structures with inductive loading. Both antennas have the same physical length, and both were loaded on the end with an inductive coil. The body of one antenna was a straight metal line, and the other antenna employed a meander-line shape. These antennas were also covered by a single-layer dielectric coating.

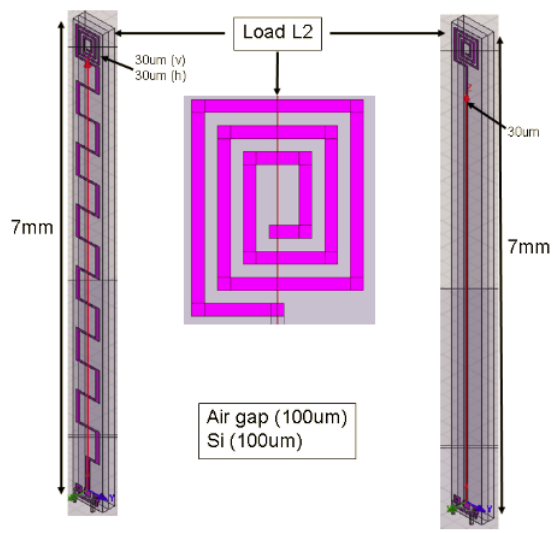


Figure 11: Antennas with inductive loading.

Table 6 shows the simulation results for these antennas. These devices exhibit performance slightly better than the previously modeled single coating 7 mm antennas.

Table 6: Simulation results for antennas with inductive loading.

#	Metal Al (3µm)	Antenna	Air Gap µm	Si thickness	Coating ε_r=4	Gain @2.4 GHz dBi	Impedance @2.4 GHz ohms
1	Meander	7 mm x 30 µm	100	100 µm 20 Ω-cm	H=8 mm D1=5 mm	-7.5	4.6-j197
2	Linear					-7.7	4.8-j260

ANTENNAS WITH BOTH CAPACITIVE AND INDUCTIVE LOADING

In this approach, the antenna is loaded with both capacitive and inductive elements. This structure of antenna is also called "slow-wave structure". Figure 12 shows the slow-wave antenna.

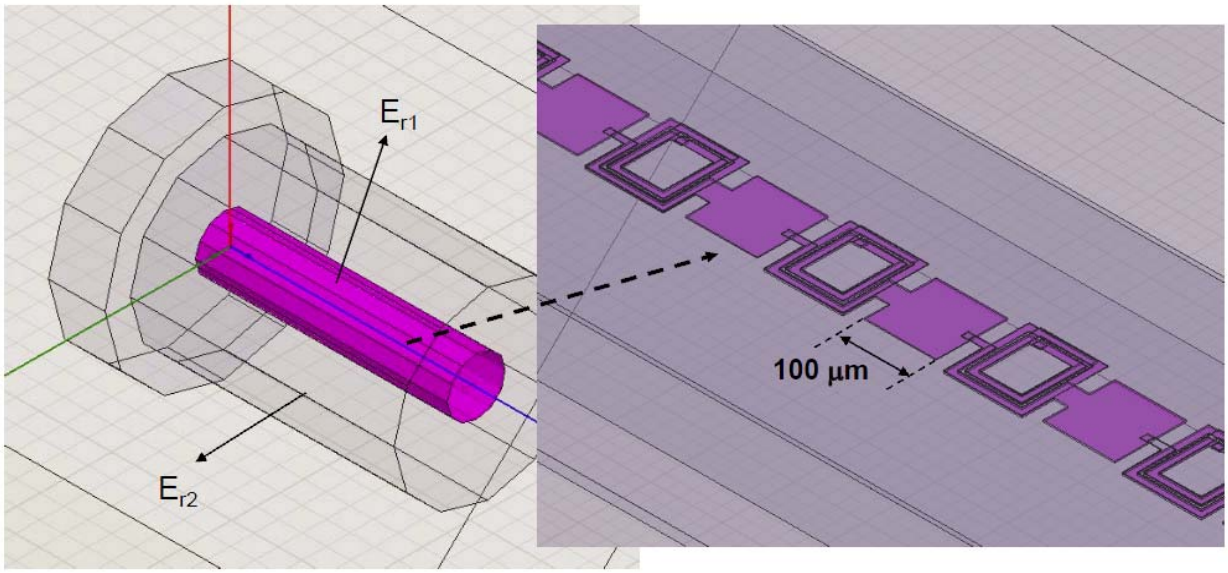


Figure 12: Slow-wave antenna.

The slow-wave antenna consists of inductive loops for the top layer and capacitive patches for the bottom layers. This provided inductive and capacitive loadings for the antenna. A ground shield was included underneath the signal pad.

Table 7 presents simulation results for some slow-wave antennas and some slow-wave antennas connected with linear antennas (for increasing the antenna length). The slow-wave structures give higher input impedance but antenna gain is lower. This is because most of the impedance comes from the loss (from silicon substrate, dielectric, antenna material and etc.).

Table 7: Simulation results at 2.4 GHz for slow-wave antennas.

#	Metal	Antenna	Air Gap µm	Silicon 20 Ω-cm µm	Coating 1 ε_r=10 mm	Coating 2 ε_r=4 mm	Gain dBi	Impedance ohms
1	Cu 2 µm Al 1 µm	6 mm slow wave	100	100	D1=2.4	H=10 D2=7.8	-18.26	170+j194
2		6 mm slow wave + 4 mm linear				H=11.5 D2=7.8	-16.01	191-j424
3	PEC 2 µm PEC 1 µm				None	None	-16.14	14.22-j294.53
4					D1=2.4	H=11.5 D2=7.8	-6.31	9.88+j189.87

ON-CHIP ANTENNAS OPERATING AT 60 GHZ

For some µNode applications it may be desirable to operate at higher frequencies where the smaller wavelengths are conducive to forming on-chip antenna arrays. For such cases the patch antenna is a popular choice because of design flexibility. Microstrip patch antenna can be planar or conformal, can be fed in numerous configurations, and are compact. Figure 13 shows the cross section of a microstrip patch antenna formed on the top surface of an integrated circuit. Typically the "patch" is a rectangular metal pattern located over a larger ground plane metal separated by a dielectric material.

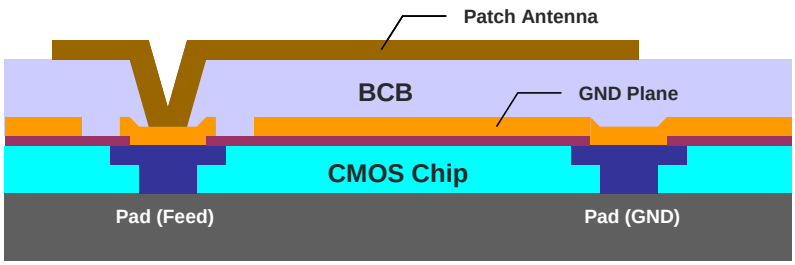
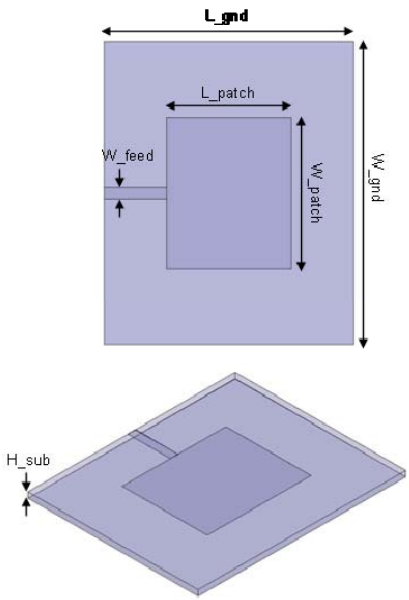


Figure 13: Cross section of patch antenna formed on a CMOS chip

Antenna performance depends primarily on the dielectric constant and thickness of the material between the patch and the ground plane. Thicker and lower index materials are needed to achieve a higher antenna gain and bandwidth. A Dow Chemical polymer, BCB (Benzo-cyclo-butene), was used for the design investigated here due to its low dielectric constant (ε_r = 2.65) and low loss (tanδ = 0.0008). BCB resins have been used in a wide variety of electronic applications, including silicon and compound semiconductor passivation, interlayer dielectric, flat panel display, IC packaging, integrated passives, MEMS, wafer bonding and 3D integration, and optoelectronic components.

This investigation was conducted in two stages. First a preliminary design was carried out to explore the approach, and then that data was used to establish a final design. Figure 14 summarized the preliminary design of the patch antenna.

The patch geometry was initially configured using PCAAD, and then the antenna parameters were simulated using a 3D electromagnetic simulator (HFSS™). Five different substrate thicknesses (H_sub) were simulated. As shown in the table included in Figure 14, antenna efficiency dropped significantly as the thickness was reduced. A thickness of 50 µm was selected for the final antenna design considering radiation efficiency and fabrication feasibility. An aluminum thickness of 2 µm was used for both the patch and the ground plane metal.



Target fr [GHz]	60				
L_patch [mm]	1.85				
W_patch [mm]	1.52				
Cond_patch [S]	3.8E+07 (Aluminium)				
t_patch [um]	2				
L_gnd [mm]	3.7				
W_gnd [mm]	3.04				
Er_sub	2.56 (BCB)				
Loss tan_sub	0.002 (assumed)				
H_sub [um]	25	50	75	100	200
Directivity [dB]	6.86	6.82	6.78	7.00	6.58
Gain [dB]	1.77	4.68	5.55	6.13	6.16
Eff [%]	31.0	61.1	71.3	81.7	89.9
fr [GHz]	59	58.5	58	57	56
Rin @fr [ohm]	82	178	200	213	220

Figure 14: Preliminary design of the on-chip patch antenna

Figure 15, Table 8 and Figure 16, respectively present the final design parameters and simulation results. An inset microstrip was designed to match the antenna to the feed-line impedance (50 ohms). The simulated resonant frequency of the antenna was 59.6 GHz. The antenna minimum return loss was 32 dB, and the impedance bandwidth was −10 dB of 0.9 GHz (1.5 %). The simulated maximum gain was 4.9 dBi at 60 GHz with a radiation efficiency of 62 %.

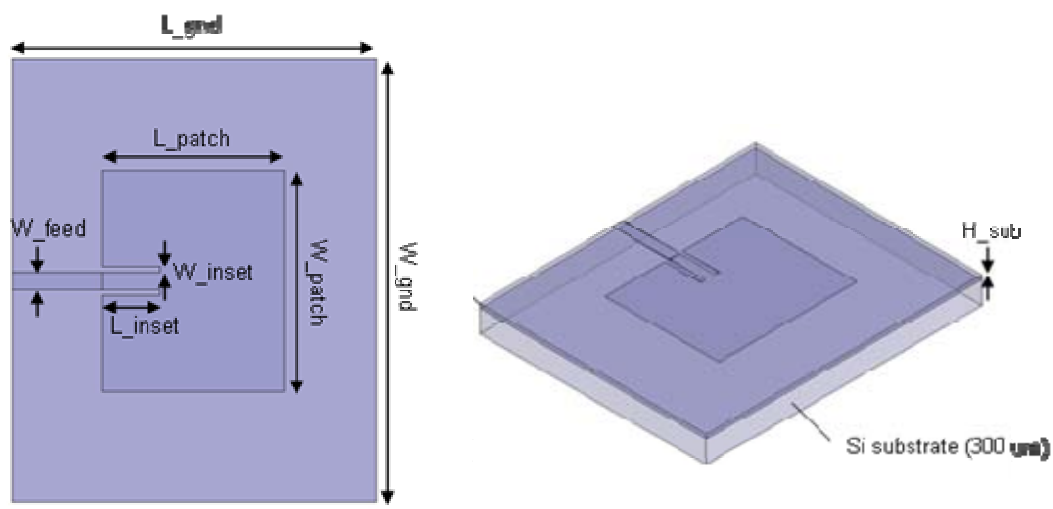


Figure 15: Final design of the on-chip patch antenna.

Table 8: Patch antenna final design parameter and simulation results

fr (GHz	59.6	W feed (µm)	138
L patch (mm)	1.85	L ground (mm)	3.7
W patch (mm)	1.52	W ground (mm)	3.04
Conductor patch (S)	3.8E7 Al	Conductor ground (S)	3.8E7 Al
t patch (µm)	2	t ground (µm)	2
εr substrate BCB	2.56	Directivity (dB)	6.96
Loss tan substrate	0.002	Gain (dB)	4.90
H substrate (µm)	50	Efficiency (%)	62
L inset (µm)	480	S11 at fr (dB)	-32
W inset (µm)	50	BW at -10 dB (GHz)	0.9 (1.5%)

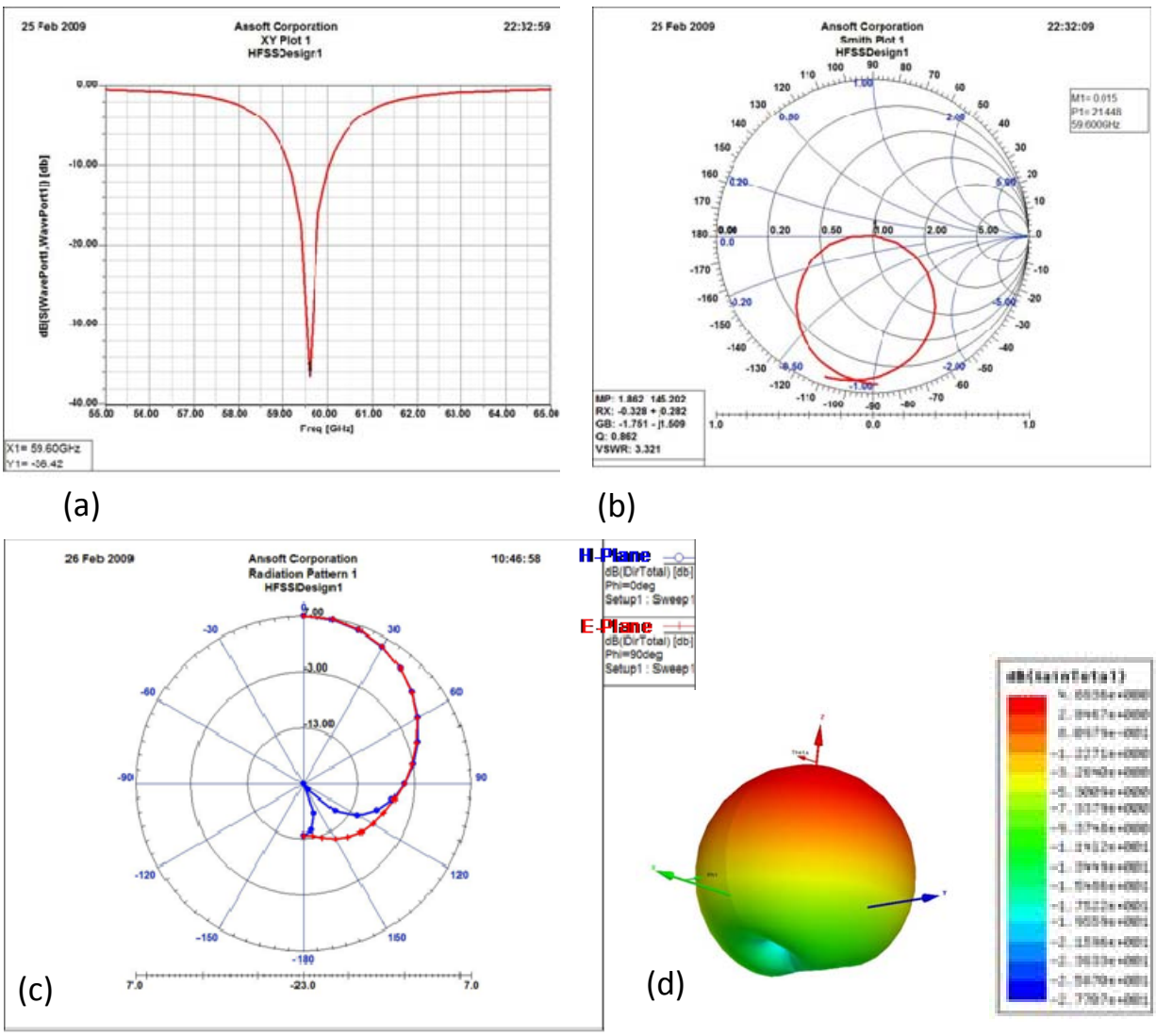


Figure 16: HFSS simulation results: (a) Input return loss (S_{11}) (b) Smith chart (S_{11}) (c) Radiation patterns (d) 3D polar radiation plot.

The patch antenna can be fabricated above an integrated circuit chip using the backend process flow illustrated in Figure 17. The idea is to first deposit the ground plane contact metal, then deposit the BCB layer, and finally deposit the top level patch and drive metal. These low temperature operations should not disturb the underlying integrated circuitry.

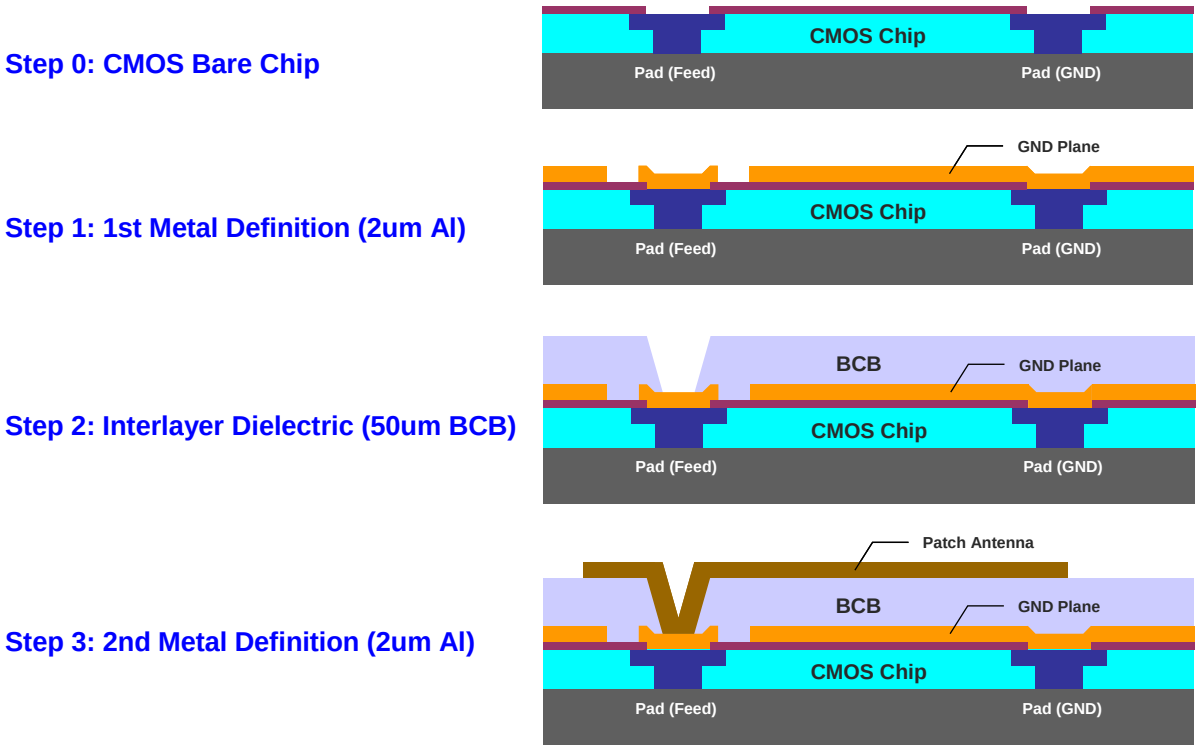


Figure 17: Back-end process flow for the on-chip patch antenna

ON-CHIP FREQUENCY REFERENCE

CRYSTAL AVOIDANCE CHALLENGE

A typical state of the art highly integrated radio contains about 9 to 15 components in addition to a power source and printed circuit board. For instance, a 2.4-GHz ZigBee radio requires a crystal, two crystal load capacitors, three decoupling capacitors, an RF IC, an antenna, and a balun to implement the system. Any off-chip component will add cost and packaging issues that will compromise the µNode cost and size objectives. The use of an off-chip crystal to stabilize the frequency reference is particularly objectionable in that the crystal itself will have to be protected against vibration, shock and temperature variation. The preferred µNode design approach is to devise a means for relaxation of the stability requirement for the reference frequency such that it is practical to use an on-chip non-crystal stabilized oscillator. Working with Motorola Labs, UFL has established and verified a technique called "differential chip detection" (DCD) that accomplishes this. The initial DCD based µNode design was carried out at 24 GHz, and for the purposes of this research it was redefined for use at 5.2 GHz or 2.4 GHz.

DIFFERENTIAL CHIP DETECTION APPROACH

Differential detection can be used as a low-complexity alternative to coherent detection in systems that can withstand a modest sensitivity penalty [COUC93]. Differential phase shift keying (DPSK) and differential quadrature phase shift keying (DQPSK) are common examples in which a delayed version of the previous symbol is used as a phase reference for demodulating the present symbol. These methods can tolerate a small amount of phase drift between adjacent symbols, and frequency offsets are limited to a fraction of the symbol rate. For direct sequence spread spectrum (DSSS) systems, performing differential detection at the chip level instead of symbol level extends the frequency offset tolerance to a fraction of the chip rate [GORD04] [COLA02] [SHI02] [CAVA97]. For typical DSSS systems, with processing gains of 10 to 30 dB, this represents a relaxation of one or more orders of magnitude in the frequency stability requirements of the transmitter and receiver.

The basic processing steps used in differential chip detection are illustrated in Figure 18. An input chip at time index k includes unwanted phase noise θ_k and frequency offset ω_ε. The differential detector multiplies the present chip by the conjugate of the previous chip, thereby converting the frequency offset to a phase term, $\omega_\varepsilon T_c$ and producing a differential phase noise term $\Delta\theta_k = \theta_k - \theta_{k-1}$. If the PN sequences representing each symbol are differentially encoded prior to transmission, then the differential chip detection, $c_k c_{k-1}^*$, will produce the desired PN sequence values at its output.

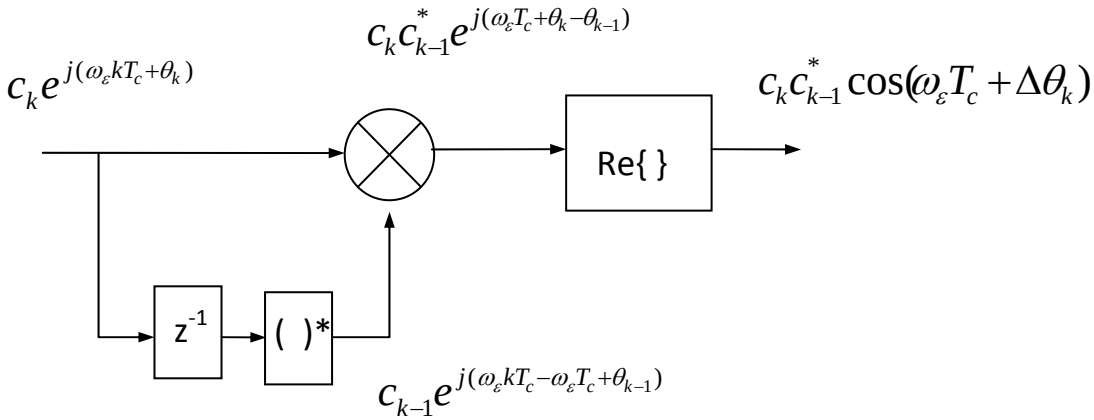


Figure 18: Differential chip detection block diagram

BASELINE PHY OVERVIEW

The baseline PHY uses a DSSS technique in which each data symbol is represented by one of 16 different PN sequences. The PN sequences are selected to be approximately orthogonal, and as a result, this technique can be viewed as 16-ary orthogonal modulation. The DCD bits corresponding to the selected PN sequence are modulated onto the carrier using Offset QPSK (O-QPSK) with half-sine pulse shaping, as shown in Figure 19.

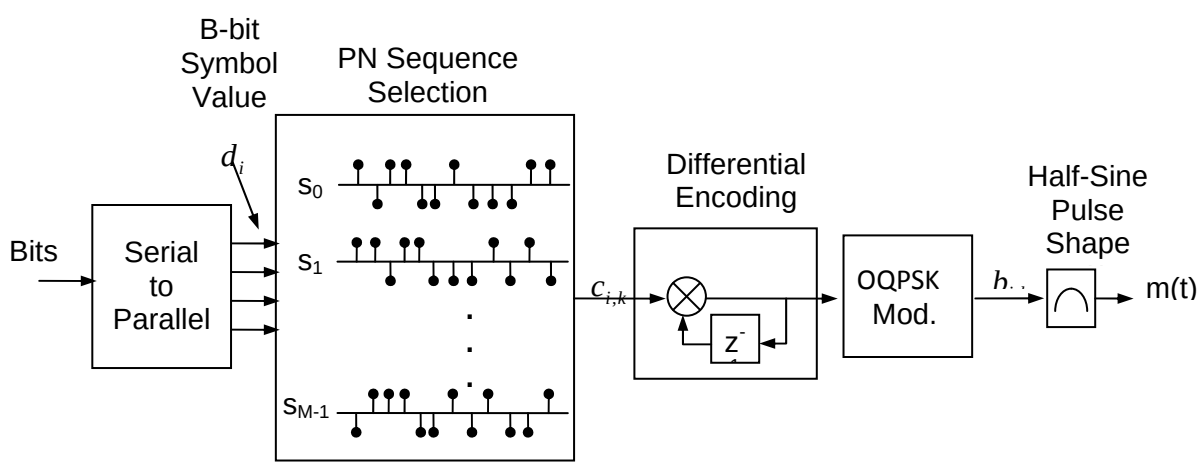


Figure 19: Block diagram for M-ary quasi-orthogonal modulator.

The modulation format extends the use of DCD to 16-ary orthogonal signaling, offering improved detector performance at the expense of increased demodulator complexity. For general M-ary orthogonal signaling, a group of B information bits is used to select one of $M = 2^B$ orthogonal waveforms for transmission during a symbol period. The M=16 orthogonal waveforms are actually different PN sequences, making it possible to apply DCD during demodulation. The set of sequences $\{s_0, s_1, \ldots, s_{M-1}\}$ comprising the M-ary symbol alphabet consists of M cyclic shifts of an m-sequence. Since m-sequences

are known to have good autocorrelation properties, the resulting set of sequences will have good cross-correlation properties (nearly orthogonal), as well as good autocorrelation properties for each individual symbol.

BASEBAND RECEIVER ARCHITECTURE

The µNode digital baseband receiver is comprised of a differential chip decoder (DCD), inner and outer correlators, a preamble detector, symbol demodulator, start of frame delimiter (SFD) detector and received signal strength indicator (RSSI) calculation block. Figure 20 presents the functional block diagram for the receiver.

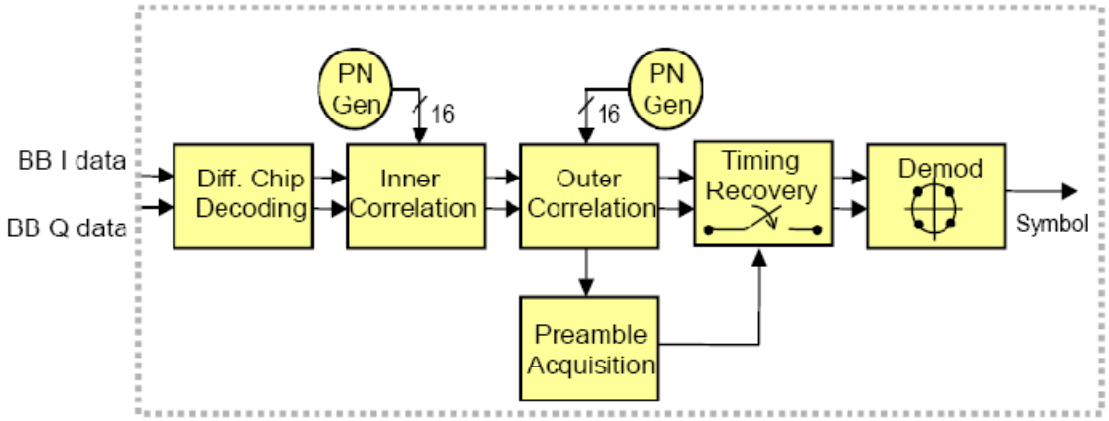


Figure 20: Block diagram for the µNode Receiver

The receiver implements a direct-sequence spread-spectrum transceiver with differential chip detection. The modulation format is a constant envelop O-QPSK with half-sine pulse shaping (or MSK). The data rate is 100Kbp/s equivalent to a symbol rate of 25Ksym/sec. Since the each symbol is spread by a factor of 256 the resulting chip rate is 6.4MChips/sec.

Each packet consists of preambles, start of frame delimiter (SFD), number-of-payload (NOP), variable payload, and an optional cyclic redundancy check (CRC). The preambles are used for symbol synchronization and AGC control, and frame synchronization. The differential chip detector operates on 5-bit I and Q samples from the A/D channels to remove phase offsets between transmitter and receiver, while mitigating frequency offsets as well as phase noise.

A cascaded hierarchical PN code structure (16 inner PN code and 16 outer PN code) is employed to implement the correlation. This structure simplifies the correlator implementation as it requires only N+M taps, as opposed to N x M taps for an arbitrary PN sequence of the same length. From the perspective of the correlator, the two main modes of the receiver are acquisition and demodulation. In acquisition mode, the preamble is being sought, and since initial timing synchronization has not yet been established, all incoming sampling phases must be correlated with all possible rotations. In demodulation mode, the preamble has been found already, so initial timing synchronization indicates symbol boundaries.

Symbol timing recovery is accomplished using information gleaned from the preamble, a sequence of symbols known a-priori by the receiver. The border between individual preamble symbols is determined and used to produce an initial estimation of symbol time slots, based on inner correlation results stored in the input memory of the outer correlator. The received signal strength indicator (RSSI) block estimates the energy in the channel by calculating the magnitude of the incoming baseband signal over a symbol period, effectively giving the average RSSI.

DIGITAL BASEBAND PROCESSING AND DESIGN

The system architecture of a µNode digital baseband processor is shown in Figure 21. It consists of two main subsystems: the transceiver and the processor. The transceiver subsystem receives and transmits signals as commanded by the processor subsystem. The processor subsystem executes the communication protocol and records application data. Commands and status are communicated between the two subsystems over an all-purpose bus, while received and transmitted symbols are transferred over a high-performance bus.

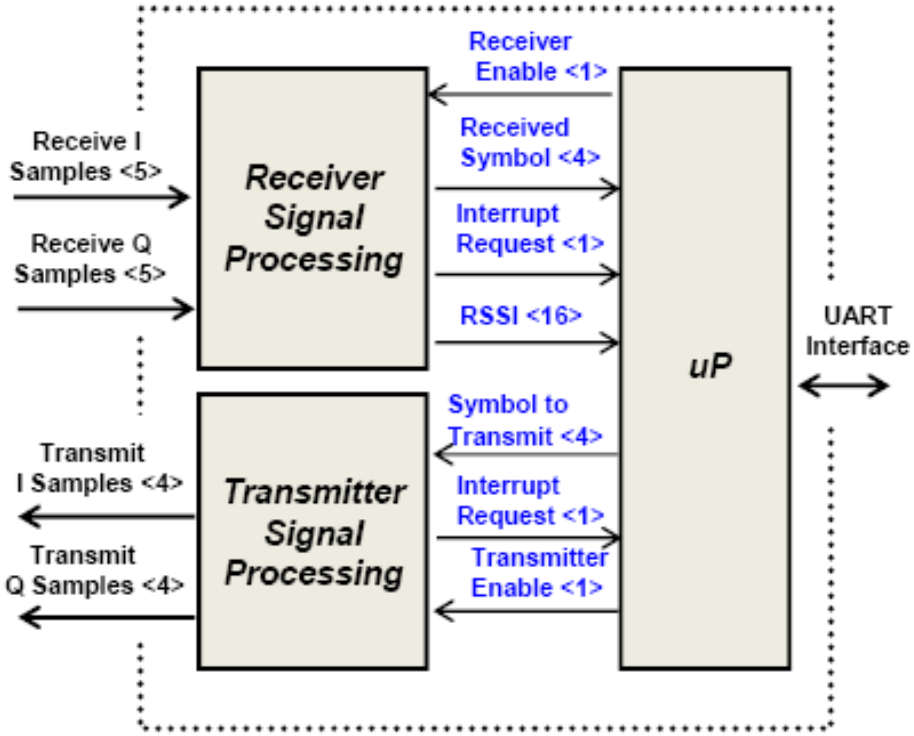


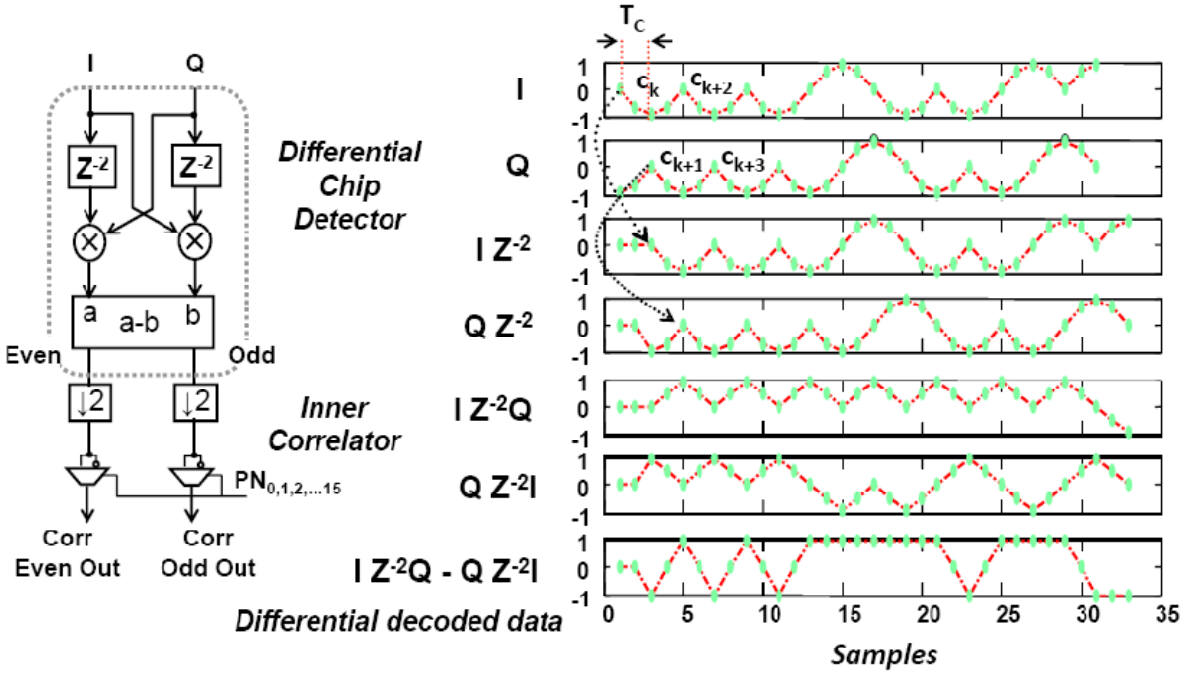
Figure 21: Block level diagram of the µNode digital signal processor

DIFFERENTIAL CHIP DECODER AND INNER CORRELATOR IMPLEMENTATION

DCD is the first block in the signal path of the receiver. The I and Q signals are received as half sine pulse shaped signals, four samples each 5 bits long represent each half sine wave. As shown in Figure 22a, these samples are multiplied by complementary samples delayed by one chip period, and then the

results are subtracted. This implements both the parallel to serial conversion and differential chip decoding simultaneously.

Since there are four samples for each half-sine pulse, the DCD outputs two samples per chip, which are even and odd phases. The even correlator selects only the even phases, while the odd correlator selects only the odd phases. The detailed waveforms of the inputs and outputs of DCD are shown in Figure 22b. Two phases, i.e. even and odd are generated for one chip c_k. The same thing happens for c_{k+1}. Due to the subtraction, if c_k is the chip, then c_{k+1} is the inverted version of the chip. Thus the resulting PN sequence is the alternately inverted version of the PN sequence from the transmitter. Table **9** shows a sample sequence encoded by the transmitter and decoded by the receiver. At the transmitter each PN sequence bit is simply XOR'ed with the previous differentially encoded bit, and a reverse of this action takes place at the receiver.



(a) (b)

Figure 22: Differential chip detection (a) functional schematic and (b) timing waveforms

A processing gain of 256 chips per symbol is implemented using a hierarchical PN sequence comprised of a 16 chip inner correlator and a 16 chip outer correlator. A block level functional diagram of the inner and outer correlators is shown in Figure 23. The inner correlator is implemented using a simple digital shift register. Since the DCD operates on 5 bit I and Q data sequences, the inner correlator is implemented with 10x16 elements. Multiplications of the PN chips by samples are implemented using multiplexers. The output of the inner correlator is fed to the outer correlator which exhibit peaks at every 16 samples as a result of the inner correlation. The outer correlator is implemented using a

memory file approach. In order to reduce the power dissipation, rather than shifting each register every clock cycle, individual registers are individually enabled by a pointer to store an incoming sample.

Table 9: Differential Chip Encoding/Decoding

Transmitter								Receiver			
Positive PN 16				Negative PN 16				Recovered Positive PN 16			
PN	ENCODED	Q	I	PN	ENCODED	Q	I	Q	I	DCD	PN
0	0	0		1	1	1		0		0	0
0	0		0	1	0		0		0	0	0
1	1	1		0	0	0		1		1	1
1	0		0	0	0		0		0	1	1
1	1	1		0	0	0		1		1	1
0	1		1	1	1		1		1	0	0
0	1	1		1	0	0		1		0	0
1	0		0	0	0		0		0	1	1
0	0	0		1	1	1		0		0	0
0	0		0	1	0		0		0	0	0
1	1	1		0	0	0		1		1	1
0	1		1	1	1		1		1	0	0
1	0	0		0	1	1		0		1	1
1	1		1	0	1		1		1	1	1
0	1	1		1	0	0		1		0	0
1	0		0	0	0		0		0	1	1

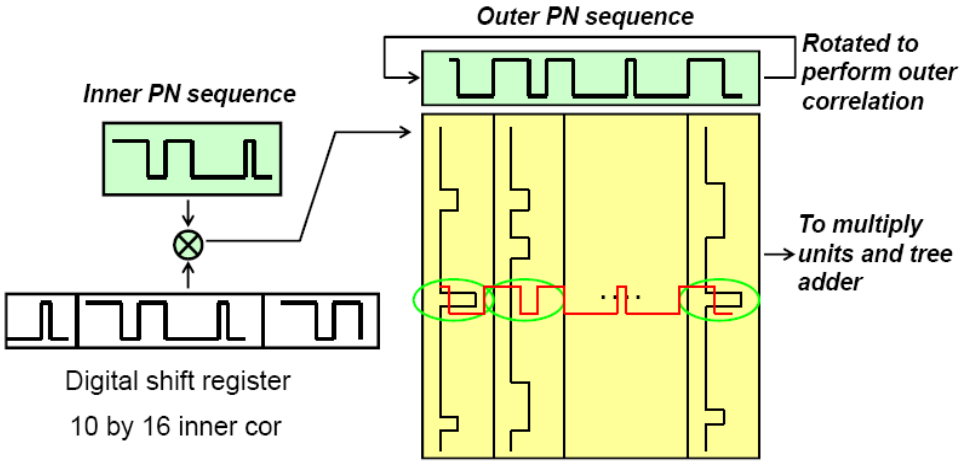


Figure 23: Correlator Peak detection

Figure 24 shows the memory architecture for the outer correlator. It consists of 16 two port SRAMs each 32X8 bits. The first 16 rows form Bank 0 and the remaining rows act as Bank1. The outer correlator operates in two modes: acquisition mode and demodulation mode. During acquisition mode, the start of a preamble in a packet is sought. All timing phases (rows) are correlated with all 16 circular shifts of a 16-chip preamble PN sequence as shown in Figure 24. A counter keeps track of the instant in time when the peak occurs.

Once the symbol synchronization is achieved using the preamble, the start of following symbols is located. A ping-pong addressing scheme is used to fill the register file to relax the timing constraints. While one bank is filling with incoming samples, demodulation is performed in the other bank as was shown in Figure 23. Dotted arrows indicate the way the samples are filling the memory. During the demodulation mode, only five phases (-2, -1, 0, +1, +2) are needed for time-drift tracking. While the first five phases in Bank 0 is enabled for write, the dotted area in Bank 1 is used for correlation. Even though the preamble has been acquired, symbols can drift depending on the channel conditions. By using only 5 phases, overall power consumption can be reduced, since this avoids writing and correlating all the phases during demodulation.

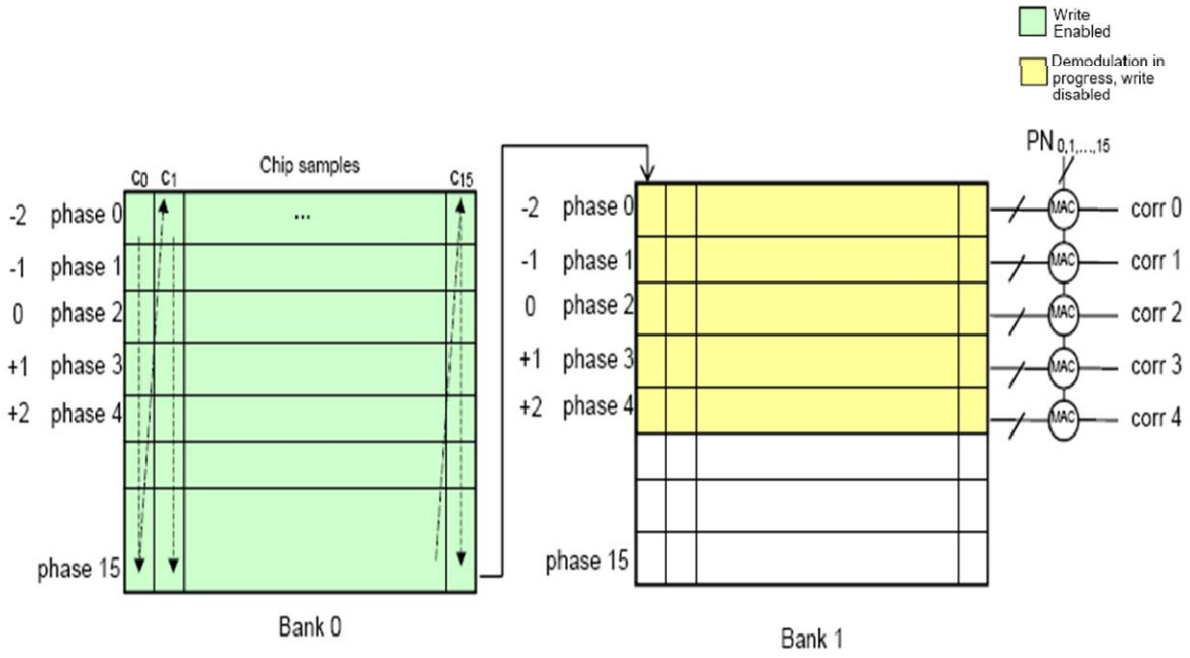


Figure 24: Memory Architecture of Outer Correlator

TRANSMITTER ARCHITECTURE

As mentioned earlier the baseline PHY uses a DSSS technique in which each data symbol is represented by 4 data bits. Each 4 bit data symbol had to be mapped to a 256 bit long PN sequence. This is accomplished by using a hierarchical PN code structure as mentioned earlier. As can be seen in Figure 25 based on the incoming symbol bits one out of 16 PN sequences is selected as the outer PN code. The PN sequences are selected to be approximately orthogonal, and as a result, this technique can be viewed as 16-ary orthogonal modulation. Based on each bit of the outer PN code the sign of inner PN sequence is decided. A set of cyclically shifted m-sequences is used to form the PN sequences. The symbol sequences are concatenated and passed to the differential chip encoder. The differentially encoded bits corresponding to the symbol are modulated onto the carrier using Offset QPSK (O-QPSK) with half-sine pulse shaping.

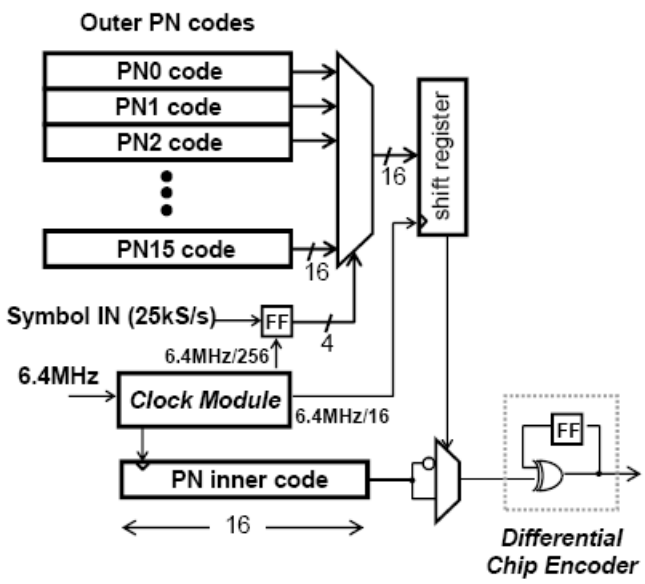


Figure 25: Transmitter Block

ON-CHIP REFERENCE OSCILLATOR APPROACH

Differential chip detection (DCD) relaxed the requirements for reference oscillator stability to the point that crystal stabilization is not required. For the specific DCD parameters established for this research a stability of ±100 ppm over the anticipated temperature and voltage operating ranges was more than adequate. From the point of view of a µNode assembly operating in the field, the battery characteristics will define the allowable temperature range. For the purposes of oscillator design, that limitation was ignored, and the circuit was planned for operation over the entire military temperature range. The design goal was to achieve a frequency stability of ±100 ppm over the -55°C to 125°C temperature range. To assure accuracy of the reference frequency, the goal was to incorporate circuitry into the oscillator subsystem that facilitates a onetime initial calibration followed by automatic compensation for temperature changes during normal operation.

Working under a subcontract, Kairos Microsystems Corporation performed the design of a suitable oscillator based on a 130 nm CMOS process. Simulations confirmed that the circuit achieved better than the ±100 ppm stability required. For the purposes of the power minimization study, the current and power estimates obtained from the 130 nm design were scaled to estimate the current and power for a 65 nm CMOS design.

POWER MINIMIZATION

The chosen research vehicle, a µNode, when used for a semi-covert type of military application requires that the node size be so small as to be non-obvious to the casual observer. Size will be driven by the battery dimensions since the µNode chip itself will be only a few millimeters on a side. In turn, the battery size will be driven by the power and energy requirements for the mission. Thus, minimum power dissipation in the µNode functional blocks is an essential requirement.

In a seedling effort with severe limits on time and funding, it was not appropriate to do the complete analyses required to fully determine power dissipation in all the individual blocks. Nor was it appropriate to explore various switching power management schemes. The effort was focused on setting workable dissipation goals for the functional parts of the µNode as implied by the visualized mission, and then examining via simulation whether it is feasible to meet those goals.

For the purposes of this research, it was assumed that the µNode will be battery powered. As shown in Figure 26, two 85% efficient regulators were assumed to be used for developing the voltages needed for the RF mixed signal circuits and the digital circuits.

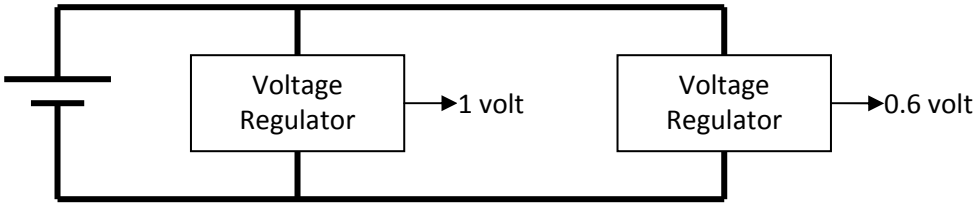


Figure 26: Power system for µNode

POWER SOURCE ISSUES

The µNode power source must meet both energy density and peak current demands. The best present day battery technology for energy density is the zinc air option, but a zinc air battery has very limited peak current capability. Large capacitors would have to be added to provide peak currents, and the added cost and volume would make it difficult to meet the µNode cost and size objectives. Preliminary µNode designs have been carried out assuming a CR1025 lithium/manganese dioxide battery with an energy capacity of 30 mAh. It has the capability of sourcing peak pulsed current of 15 mA. At 0.1% duty cycle and peak power consumption of 5 mW, this small sealed battery has sufficient energy for an operating lifetime of one year which is sufficient for most µNode applications. In addition, this coin battery is small enough to allow realization of the desired µNode marble configuration.

BLOCK SHUT-DOWN STRATEGY

Establishment of an over-all power management strategy and specific scheme will play an important part in the eventual final design of a µNode device. This fact is mentioned here to make clear that this aspect of power management is not being overlooked. It is, however, not going to be treated in any

detail during this seedling effort. The system plan calls for being operational in transmit or receive modes about 0.1% of the time. For 99.9% of the time the system will be sleeping. Design of an appropriate activation scheme is to some extent mission dependent, and thus is a development task (as opposed to being a research task). It is anticipated that leakage current and the power associated with standby circuits will constitute the power dissipation during the sleep.

POWER DISSIPATION BASELINE TARGETS AND SIMULATION RESULTS

The power minimization study considered several alternative overall architectures for implementing the µNode functions, and examined alternative ways to design individual functional blocks within those architectures. Table 10 captures both the power design targets established for individual functional blocks and the simulation results for the most representative version of the functional block.

The goal for the total µNode chip operating power was 5 mW. Operating power means the power during a receive operation, or the power during a transmit operation. Receive and transmit do not occur at the same time. This objective was intended to provide some indication of what the peak demands would be on a battery. Receive and transmit will occur with a low duty cycle (0.1 %), so the average power will be low. Estimation of the various standby type modes that would determine the average power requires a more detailed design, and was beyond the scope of this seedling research.

The simulations confirmed that it is reasonable to assume that achievement of operating power levels near 5 mW is feasible. A metric that is often mentioned for transceivers is the communication energy. Including both transmit and receive energy for 100 kbps, the communication energy was projected to be 100 nJ/bit.

Table 10: Summary of µNode operating power dissipation

Rx Blocks	Supply (V)	Power (mW)	
		Target	Simulation
Mixer	1	0	0.88
LNA	1	0.7	0.46
Multiphase	1	0.4	0.18
VGA/LPF	1	0.5	0.52
PLL/VCO	1	2.5	2.41
ADC	1	0.3	0.3
Bias	1	0.1	-
Subtotal	1	4.5	4.75
Digital	0.6	0.8	0.46
Total		5.0	5.21

Tx Blocks	Supply (V)	Power (mW)	
		Target	Simulation
PA	1	1	1.48
Modulator	1	0.5	0.49
Multiphase	1	0.4	0.18
			-
PLL/VCO	1	2.5	2.41
			-
Bias	1	0.1	-
Subtotal	1	4.5	4.56
Digital	0.6	0.8	0.46
Total		5.0	5.02

DIFFERENTIAL VERSUS SINGLE ENDED ARCHITECTURE

Obviously the choice of circuit architecture for the µNode has a significant impact on the power dissipation. The two major alternatives considered were a differential front end versus a single ended front end. Figure 27 was an example of a differential RF approach. Here the antenna interface was simplified by the use of a transformer balun. In the receiver, the first functional block was a passive RF mixer rather than a low noise amplifier (LNA). This strategy minimized the impact of the power hungry high frequency circuit.

If a double balanced mixer were used to accommodate the differential input, then eight local oscillator (LO) amplifiers would be required to form the I & Q paths. The current consumption per amplifier would be about 200 to 300 µA to provide sufficient drive in a 2.4 GHz design. Thus, the overall current consumption for just the buffer amplifiers in the receiver (RX) path amount to 1mA. In addition, two more buffer amplifiers consuming 200 to 300 µA were required to drive the differential power amplifier (PA) in the transmitter (TX) path. Considering the entire differential transceiver (TRX), the buffer amplifiers would consume 1.5 mA, or 1.5 mW from the 1 V supply voltage. With 30% of the target µNode power consumption of 5 mW being lost in buffer amplifiers, it was concluded that the differential RF front-end architecture was not an appropriate approach.

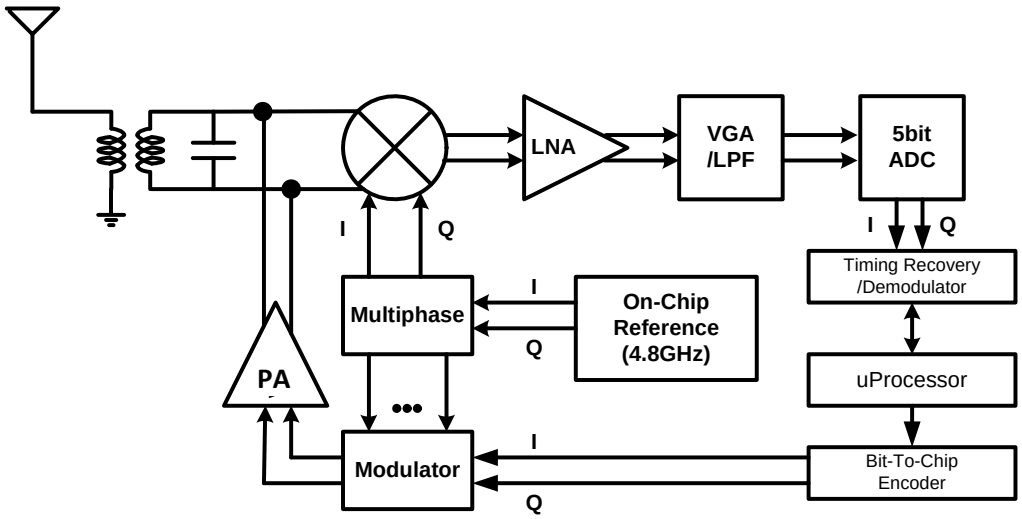


Figure 27: Differential RF front-end µNode

In the single ended architecture shown in Figure 28, an antenna connects to the front-end through an LC matching circuit. A single balanced mixer and a single-ended PA is used in RX and TX respectively. Thus, the number of the power consuming buffer amplifiers in RF mixers and the PA was half of that for the differential RF front-end architecture, and a considerable power reduction results.

The optimal load impedance for a sub-mW PA was in the range of 100's to 1000's of ohms, so the LC matching network must boost the antenna impedance. This network provides about 10 dB voltage gain at the RX input and reduces the noise figure burden of the LNA. This also mitigates the power requirement for the LNA.

For power efficiency the transmitter architecture employed a direct modulator. This was a considerable simplification compared to a quadrature mixer based approach. It avoided having to include a digital to analog converter (DAC) and a low pass filter (LPF).

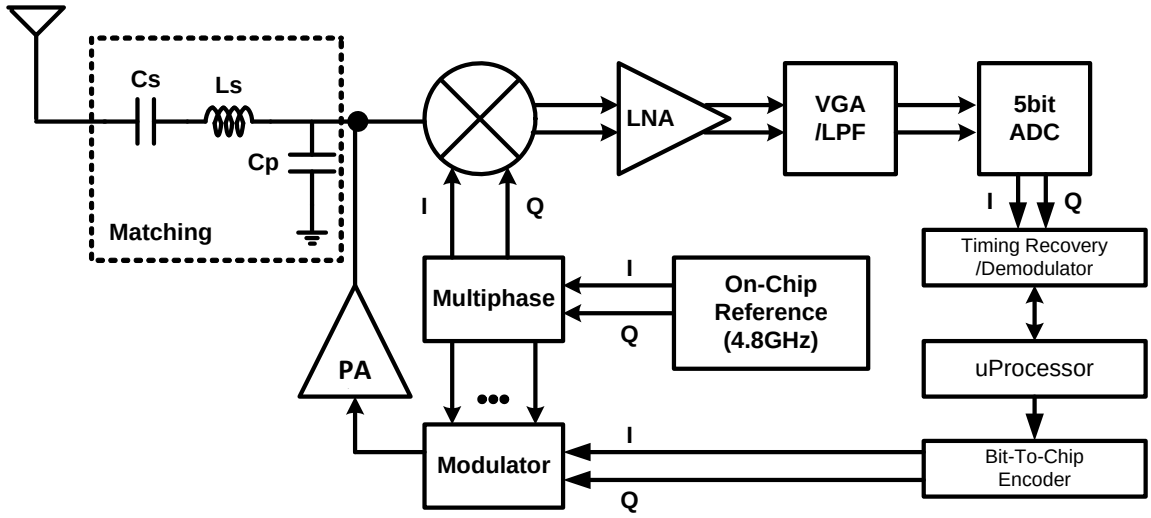


Figure 28: Single-ended RF front-end µNode

RECEIVER DESIGN

The receiver approach presented in Figure 29 down converted the 2.4 GHz RF signal directly to baseband. The antenna was interfaced to an impedance transformer that fed I and Q mixers. The baseband output was fed to an amplifier chain. In the receiver analysis, the monopole antenna was assumed to have a 50 Ω source resistance. The impedance transformer raised that impedance to 200Ω. Two balanced passive mixers formed the second stage, and generated I and Q signals using four phases from the 2.4 GHz LO. The first amplifier after the mixer has a lower noise figure than the variable gain amplifier (VGA) stages that follow. This low noise amplifier (LNA) improved the overall noise figure. The VGA consisted of nine stages where the overall gain was controlled in 6 dB steps by a 3 bit digital word. The I and Q outputs drive an analog to digital converter.

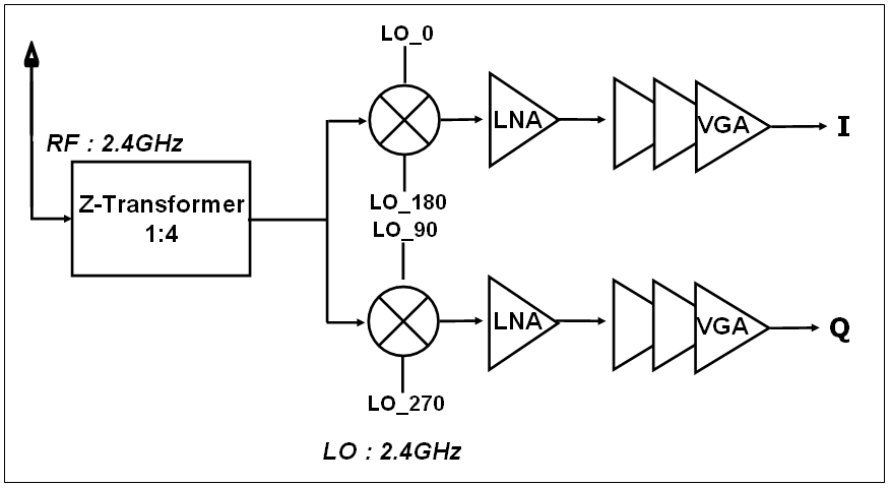


Figure 29: Receiver block diagram

One side of the front end of the receiver circuit was illustrated in Figure 30. The diagram includes the passive impedance transformer, the balanced passive down converter mixer, and a resistive feedback inverter. The two NMOS transistors forming the mixer do not consume power (excluding that of the LO driver). Because only a single mixer was required, the number of LO drivers was cut in half.

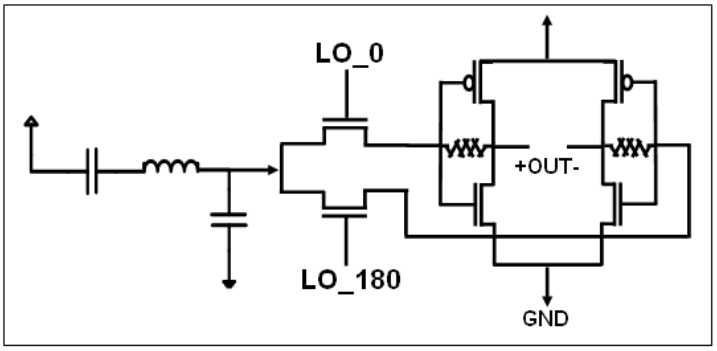


Figure 30: One half of the receiver front end circuit

The resistive feedback inverter offered some advantages. No external bias circuit was required because the feedback resistor determined the bias depending on the ratio of the NMOS and PMOS transistor widths. This arrangement avoided requiring a DC blocking capacitor that attenuates low frequency signals. .

Simulations using IBM 65 nm low power CMOS predicted 40 dB gain and 11.8 dB noise figure for the front-end circuit. Power consumption was 460 µW for the two amplifiers, and 880 µW for the four LO drivers. The total power for the front-end was 1.34mW.

Figure 31 shows the basic cell used to form the variable gain amplifier. The each cell provided 6 dB of gain, and the number of cells was activated by 3 to 8 decoder controlled by input of a 3 bit word. The R1 to R2 ratio determines the gain of the cell.

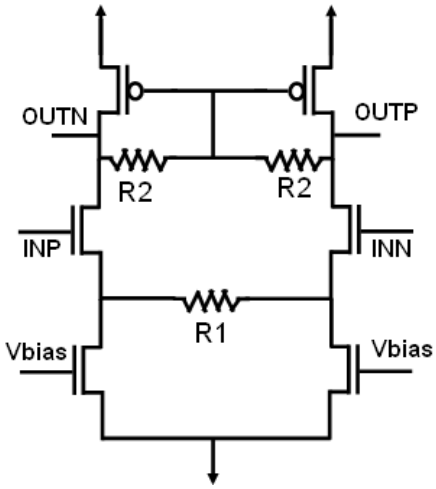


Figure 31: VGA basic cell

The first stage of the VGA has a fixed gain of 12 dB. It was followed eight basic cells that provide 6 dB each. Thus the minimum gain is 12 dB, and the total gain can be varied in 6dB increments up to the maximum of 60dB.

As summarized in Table 11, the receiver total gain was about 100dB and the noise figure was about 12dB. The power consumption was 520 µW for the two VGA's. Total power for the receiver was 1.86mW. Since none of these circuits were optimized, it was expected that further power reduction was a reasonable expectation.

Table 11: Receiver simulation results

	Z-Tr.	Mixer	LNA	VGA	Total
Gain (dB)	13.4	-3.6	30.1	60	99.9
NF (dB)	1.5	9.2	11.2	21.5	12
Power (mW)	0	0.88	0.46	0.52	1.86

TRANSMITTER DESIGN

Figure 32 shows transmitter functional blocks for the OQPSK (MSK) direct digital modulator. MSK is a constant envelope modulation scheme where phase varies linearly in the time domain.

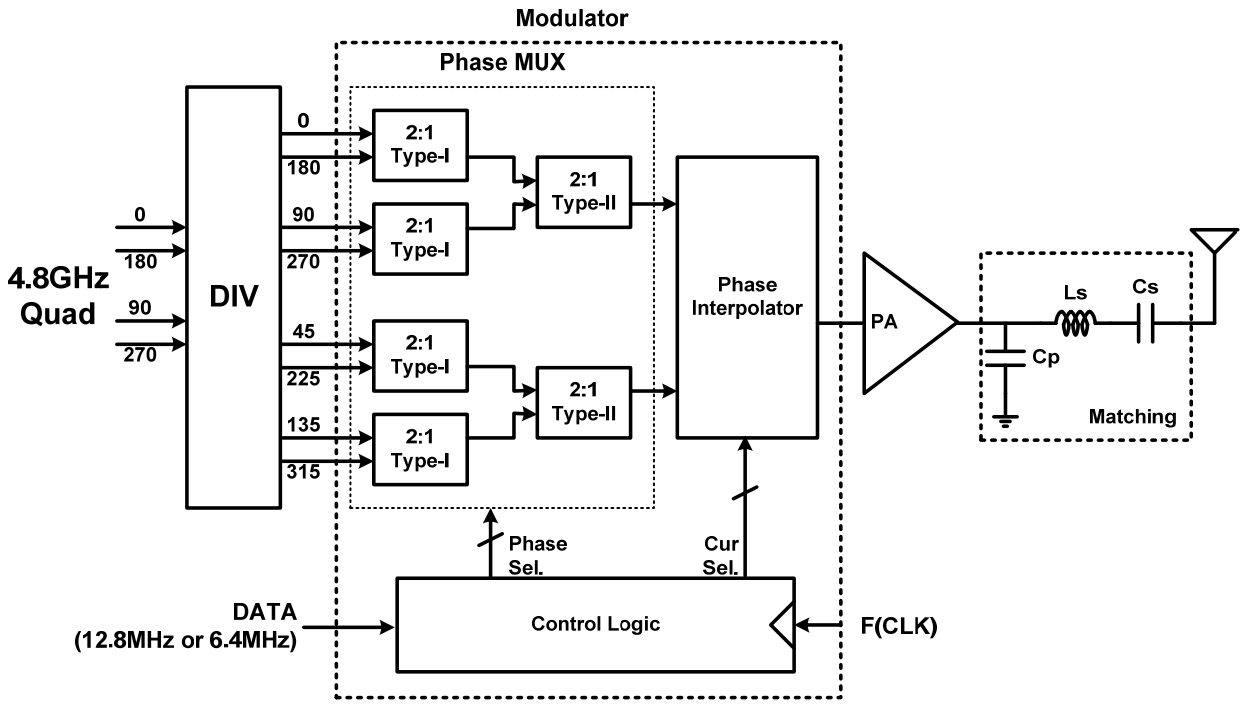


Figure 32: Transmitter OQPSK (MSK) direct digital modulator

The modulator was composed of an 8-to-2 phase multiplexer (MUX), a 2-to-1 phase interpolator, and control logic. The modulator generates a discrete (4-step) linear phase in one data period by controlling the incoming 8-phase LO signal with the phase MUX and interpolator. Then, the modulated signal is sent to the antenna via a class-D power amplifier. The frequency divider driven by a quadrature 4.8 GHz input generates an 8-phase 2.4GHz signal output.

INJECTION LOCKED 4.8 GHZ FREQUENCY DIVIDER

The structure of the frequency divider and the delay cell building block was presented in Figure 33.

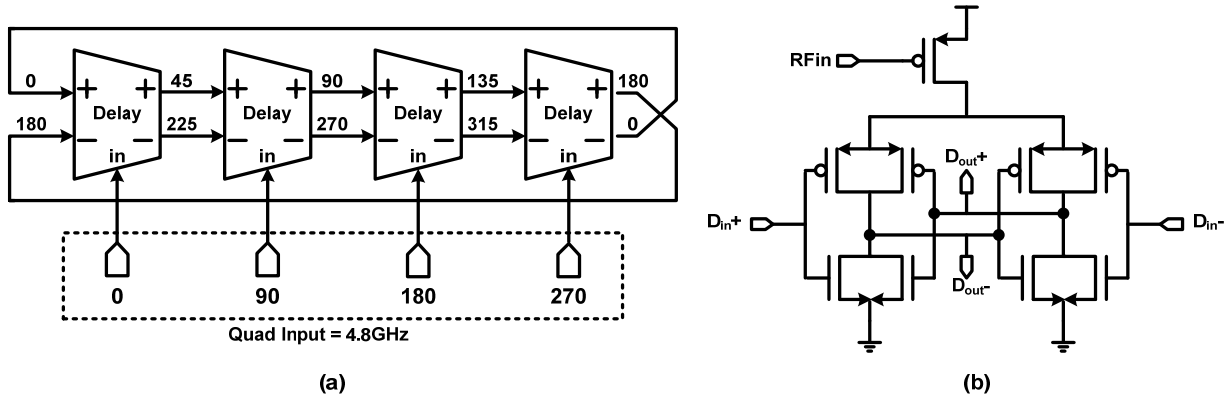


Figure 33: Injection locked frequency divider (a) and delay cell detail (b)

The 4-stages of delay cells form an 8-phase ring oscillator that oscillates around 2.4GHz. When a 4.8GHz quadrature signal was injected to the ring oscillator, the oscillator was locked to 2.4GHz. Simulation showed that the locking range of the divider was about 4GHz with a 100mV input signal.

PHASE MULTIPLEXER AND INTERPOLATOR

Figure 34 shows the Type-I and Type-II phase multiplexers. The Type-I phase multiplexer, which has a differential input, selects non-inverted or inverted phase signals using switches A and B. The Type-II phase multiplexer, which has two differential phase inputs, selects one differential input signal using switches A and B.

Figure 35 shows a phase interpolator which has two differential phase inputs. If the switch A is enabled, the S1 differential phase is selected. If the switch C is enabled, the S2 differential phase is selected. If the switch B is selected, the interpolated phase between S1 and S2 is generated at the output. Thus, the phase interpolator can generate the desired signals with eight different phases at the output.

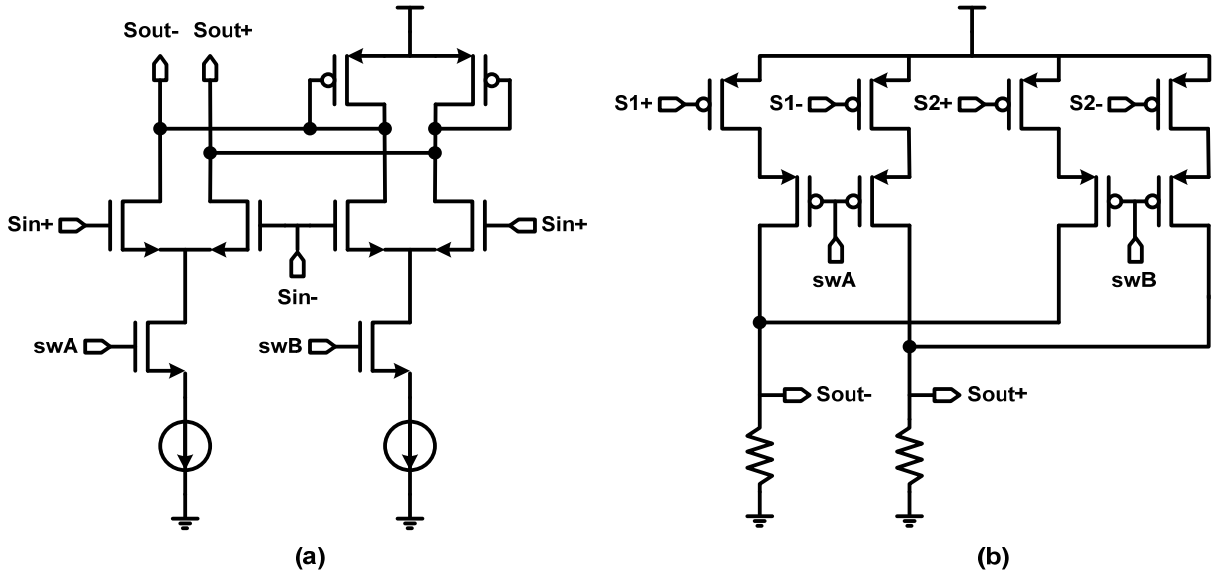


Figure 34: Phase multiplexers (a) Type-I, (b) Type-II

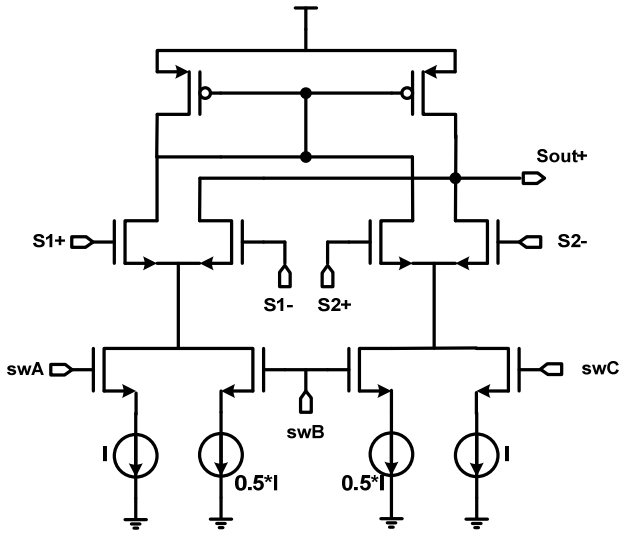


Figure 35: Phase Interpolator

CLASS-D POWER AMPLIFIER

Figure 36 shows the class-D power amplifier where the buffer amplifier was an inverter. The load impedance of PA was targeted for about 200Ω. The output power of this PA was -3dBm. The simulated drain efficiency is 42%. Power consumption is 1.15 mW for a 1 V supply voltage.

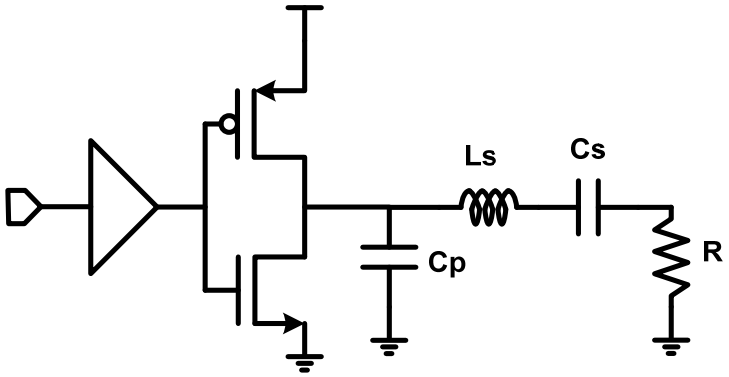


Figure 36: Class-D Power Amplifier

TRANSMITTER SIMULATION RESULTS

Transmitter simulation results were summarized in Table 12. The total power dissipation was 2.15mW for a 1V supply voltage

Table 12: Transmitter simulation results

	DIV	Phase MUX	Phase Interpol	PA	Total
Power (mW)	0.18	0.32	0.17	1.48	2.15
PA Drain Efficiency (%)	-	-	-	42	-
PA Output Power(dBm)	-	-	-	-3	-
Lock-In Range [GHz@ 100mV]	3.5	-	-	-	-

FREQUENCY REFERENCE POWER

The power consumption of the reference oscillator block was summarized in Table 13.

Table 13: Power Consumption of each block

Blocks	130 nm CMOS mW	65 nm CMOS mW
5GHz DCO	1.4	0.5
Fractional divider	3.85	1.17
Buffers + AAC	2.45	0.74
Calibration Circuits	0	0
Total	7.7	2.41

The power consumption of fractional divider was based on the assumption that the step size was 0.5. The 130 nm CMOS estimates were obtained by simulation. The 65 nm CMOS estimate was scaled from the 130 nm results.

DIGITAL POWER

The digital portion of the µNode implemented in 65 nm CMOS was estimated to dissipate about 460 µW. Considerable analysis was required to obtain this estimate because a design library was not available to the 65 nm process. The discussion below explains the how the power was investigated.

A µNode Transceiver was implemented in VHDL and synthesized to estimate the area and power dissipation. The transceiver was essentially redesigned and scaled for a processing gain of 256. The core was functionally verified. Although the goal was to implement the design in IBM's standard 65nm CMOS process, due to the lack of digital libraries, it was necessary to instead synthesize the transceiver in a 130nm CMOS process. Figure 37 shows a layout of the µNode transceiver measuring 750µm x 750µm in 130nm CMOS. The area breakdown of the core components clearly indicates the outer correlator as the dominant component of the design. Correspondingly, a significant amount of power is consumed by the memory of the outer correlator. Power estimates of the µNode operated at 6.4MHz indicate that the transceiver dissipates ~1.53mW in 130nm CMOS, and the outer correlator dissipates 73% of the total power (see Figure 38 and Table 14).

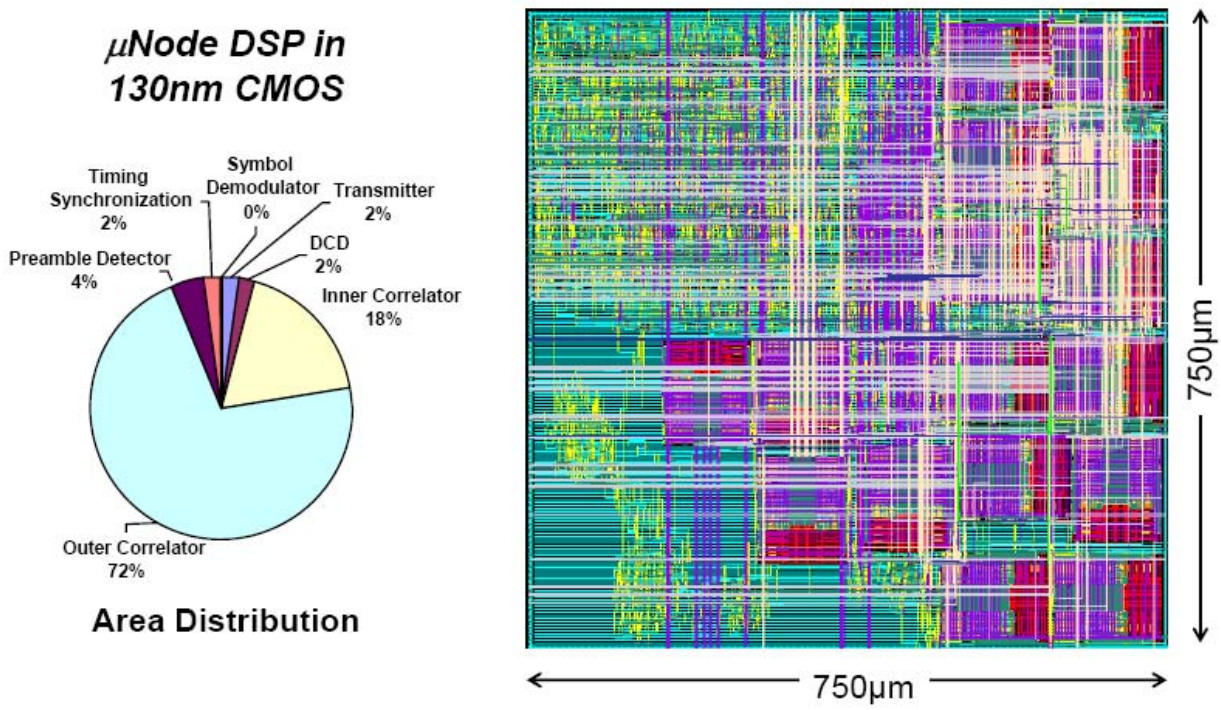


Figure 37: µNode digital signal process 130 nm CMOS layout

Table 14: Digital power estimates for 130nm CMOS process

Synthesis Power Estimate (130nm CMOS)	Power (µWatts)
Transmitter	114
Receiver	1414
DCD	35
Inner Correlator	217
Outer Correlator	1106
Preamble Detector	35
Timing/Synchronization	20
Symbol Demodulation	1
Transceiver Total	1528

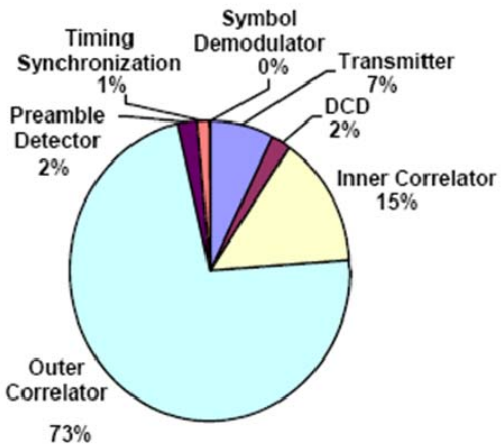


Figure 38: Digital power distribution

To estimate the power dissipation of digital baseband processor in 65nm technology, the power consumption of various logic data paths in 130nm and 65 nm technologies was compared. Table 15 shows a comparison of the power dissipation at 6.4MHz and 1V supply voltage for various digital circuits in 130nm and 65nm technologies obtained via SPICE simulations. The scaling factor for the average power ranges from 3.1x to 3.5x, and the static power ranges from 17.2x to 20.1x. Assuming an average power scaling factor of ~3.3x, this suggests that the µNode implemented in 65nm process would scale from ~1.5 mW to approximately 460 µW. A much larger scaling in leakage power is observed as the 130nm digital library employs high-speed (HS) logic whereas the target technology in 65nm uses standard devices (SD).

Table 15: Power comparison for various digital logic blocks in 130nm and 65nm

Functional Part	Type Power	130HS	65nm SD	Scaling
64 Bit FIFO	Avg. Power	6.92 µW	2.14µW	3.23X
	Static Power	199.1nW	11.53nW	17.26X
8 Stage LFSR	Avg. Power	782.3nW	249.2nW	3.14X
	Static Power	40.25nW	2.06nW	19.53X
8 Stage LFSR and a 8bit ripple carry adder and a 8:256 decoder	Avg. Power	3.526 µW	1 µW	3.52X
	Static Power	366nW	18.17nW	20.14X

In order to validate the SPICE simulations, individual transistors and inverters were characterized to estimate the capacitances, short circuit currents and leakage currents for both 130nm and 65nm logics, as shown in Table 16 though Table 19.

The results for gate capacitance per unit area indicate that the resulting equivalent oxide thickness in this 65nm process does not strictly scale to mitigate increasing gate leakage currents caused by tunneling. As a result, the expected increase in gate capacitance per unit area over two technology nodes (i.e. from 130nm to 65nm) does not double (assuming a 0.7x scaling per node). Thus, the expected power dissipation scaling factor due to dynamic power more than doubles, as indicated in Table 15.

The total capacitance scaling factor per unit micron of transistor width was ~1.5x, as indicated in Table 17 . Since the transistor width from 130nm to 65nm scales by a factor of 2x, the total switched capacitance is expected to scale by 3x. Thus, the dynamic power scales by a factor of 3x. This is slightly lower than the simulation results in Table III. This is attributed to short circuit power dissipation, which scales on average by a factor of 4x as shown in Table 16.

Table 16: Gate capacitance per unit area in 130nm and 65nm

Size	130nm HS		65nm SD	
	NMOS	PMOS	NMOS	PMOS
1u/1u	11.55 fF/ um^2	11.02 fF/ um^2	13.45 fF/ um^2	12.47 fF/ um^2
3u/3u	11.41 fF/ um^2	10.83 fF/ um^2	13.24 fF/ um^2	12.17 fF/ um^2
6u/6u	11.38 fF/ um^2	10.78 fF/ um^2	13.16 fF/ um^2	12.09 fF/ um^2
10u/10u	11.36 fF/ um^2	10.76 fF/ um^2	13.16 fF/ um^2	12 fF/ um^2

The exact breakdown between dynamic and short circuit power depends on edge rates, fanout and logic style, along with implementation details of the processor. However, it is reasonable to assume that a larger fraction of the power dissipated is due to dynamic switching capacitances. I_{OFF} leakage currents shown in Table 19 for the two technology nodes scale by roughly 21x, which appears to be in good agreement with simulated results. These results are only a rough estimate and the actual power consumption may vary slightly depending on the implementation details and accuracy of the models used herein.

Table 17: Capacitance per unit micron of transistor width in 130nm and 65nm

	130nm HS	65nm SD	Scaling Factor
Cin	1.37 fF/um	0.79 fF/um	1.73X
Cout	0.82 fF/um	0.65fF/um	1.26X
Ctotal	2.19 fF/um	1.44 fF/um	1.52X

Table 18: Short circuit current in 130nm and 65nm

	130nm HS	65nm SD	Scaling Factor
P_{scHL}	33.64	7.19	4.86
P_{scLH}	30.05	8.32	3.6
Average	31.83	7.76	4.1

Table 19: I_{OFF} and Gate leakage per unit area in 130nm and 65nm

	130nm HS	65nm SD	Scaling Factor
Ioff	14.44nA/um2	2.67nA/um2	21.2
Igon	0	51pA/um2	--

ANALOG PROCESSING ALTERNATIVES

The potential application of analog circuitry to reduce power and improve overall performance of a µNode system was investigated. Two promising alternatives were identified.

SUMMARY OF ANALOG ALTERNATIVE APPROACHES

The first approach was to use error control coding. Special classes of iterative codes including Turbo codes [BERR93], LDPC [GALL63] and Linear Block Codes [ELIA54] can help achieve very low bit error rate and be decoded using iterative decoders. These decoders are expected to provide 3dB to 6dB coding gain [VUCE00]. This coding gain allows for the RF subsystem to consume less power and relaxes the stringent constraints (size and gain) of the antenna design for the µnode project.

Typically 3dB coding gain can allow reduction in transmit power by half or a 3 dB decrease in required antenna pair gain. A 3dB increase in coding can also be used to extend the projected 20 m communication range between nodes by a factor of 1.5. Typically a BER of 10^{-5} is possible with iterative decoders [WINS05]. Possible analog implementations of these iterative decoders make them attractive for low power applications. Among many low power analog iterative decoders a (8, 4) trellis low voltage decoder consuming only 2.4µW at 0.5V for 0.18 µm technology operating at 69 kbits/s has been reported [SCHL07]. The data rate of the µnode system is 100Kbits/s which means that a decoder will be required with about the same bandwidth and technology as the decoder cited above.

The second approach targets reduced power consumption in the analog and the digital baseband systems. In the current design, the ADC in the analog baseband has to run at chip rate which is 6.4Mbps. In digital baseband the DCD decoding and DSSS “de-spreading” blocks consume most of the power. It is possible to use analog processing techniques to implement DCD decoding directly on sampled analog

values. Then using analog correlation techniques, discrete time sampled analog values can be DCD decoded. An ADC now running at a symbol rate of 25 kS/s can be used for symbol detection. The reduction in ADC operating speed is expected to save two orders of magnitude of power in the analog baseband [ONOD98]. The analog correlation also saves power by removing costly blocks such as multipliers and adders. Only the synchronization and timing control needs to be done in digital, but current research is also focusing on doing this analog [DAUW04].

Optimum power savings are achieved when both of the analog processing techniques are combined to provide system wide benefits. Incoming discrete time sampled data can be decoded using an analog DCD decoder and the samples input to an analog correlator. The output of correlators can be used as probability estimates for symbols that were transmitted and input to an analog iterative decoder.

RESEARCH PERFORMED

1. Detailed study of the digital baseband architecture to determine where analog signal processing can provide savings in power and area.
2. Study of error control codes, specifically iterative error control codes and how they can relax constraints on the RF and baseband analog subsystems.
3. A survey of published analog iterative decoders and their evaluation in terms of power and area to decide their suitability to the µNode project.
4. Simulation of a simple [5, 2, 3] linear block code using analog signal processing.
5. Evaluation of problems concerning usage of nanoscale CMOS technology for analog signal processing circuits.
6. Research on possible ways to implement differential chip detection and direct sequence spread spectrum correlation in analog to save power.

ITERATIVE DECODERS

Error control codes are used to reduce or effectively eliminate the occurrence of errors in information that is transmitted over a communication channel. Iterative decoders are a type of special error control codes which give performance close to the theoretical Shannon capacity limit. The Shannon capacity limit has long been known as a bound on the performance of error control systems [SHAN48]. Turbo Codes [BERR93], Low Density Parity Check Codes [GALL63] and Block Product Codes [ELIA54] can approach this limit on Gaussian channels. However, to achieve this high performance with these codes is still a challenging problem. This is accomplished through iterated estimation of the transmitted message. First, an entire block of data is received and decoded. Then the same block is decoded again and again achieving improved results each time. [HAGE96].

The benefit of the error control system is that it allows for error free transmission, thus relaxing constraints of signal energies, antenna performance and/or extended range. However owing to their complexity, iterative decoders can quickly outweigh these advantages by consuming more power and area when implemented in hardware with digital circuits. To overcome this limitation analog iterative decoders were proposed [WINS05]. Implementation of iteration is natural to the analog domain as it

requires only the natural settling time of the circuit. Moreover, the analog circuits work on discrete-time sampled analog values and produce a digital output in the end, eliminating the need of a costly, high speed and power hungry analog to digital converter. Figure 39 shows how an analog decoder would replace a digital decoder in a communication system.

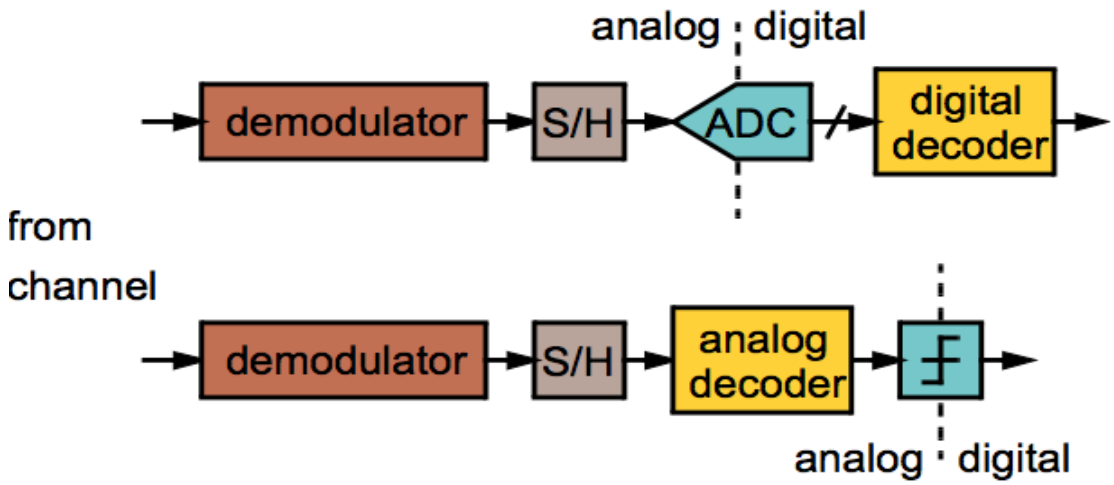


Figure 39: Replacing digital decoder with analog decoder

For the µnode system, the position of the analog decoder can be decided in two ways. If the existing decoding for Direct Sequence Spread Spectrum (DSSS) with Differential Chip Detection (DCD) is kept as it is in the digital domain, then the digital values generated by the inner correlater can be converted to soft current inputs to an analog decoder using a standard current steering digital to analog converter.

However a higher impact option is to implement the DSSS and DCD in the analog domain and use the analog values as the input to the analog decoder. Table 20 and Table 21 list the reported and fabricated results for digital and analog iterative decoders respectively. This list has been created after a literature survey and most of the references can be found in [WINS05] and [SCHL07]. More detailed references of earlier work for analog iterative coders can be found in [WINS05]. The analog iterative coders listed below are fabricated in a regular CMOS process. Many other analog iterative decoders in BiCMOS and SiGe process using BJT's have also been successfully reported but those decoders are not considered here since the µNode is based on the 65nm IBM CMOS process.

Table 20: Fabricated, measured and reported digital iterative decoders

Code	Process	Power	Throughput	Energy/Bit
512 LDPC code	0.16 µm, 1.5V	630 mw	500 Mbits	1.26nJ/b
3GPP Viterbi Turbo	0.18 µm, 1.8V	306 mW	2.048 Mbits	122nJ/b
PCCC Turbo	0.8 µm, 5V	1.6 W per it	40 Mbits	160 nJ/b

Table 21: Fabricated, measured and reported analog iterative decoders

Code	Process	Power	Throughput	Energy/Bit
Small turbo code	0.35µm	185mW 3.3V	13.3 Mbits	13.9nJ/b
(8,4) Hamming trellis	0.5µm	45mW 3.3V	1 Mbits	45nJ/b (core + IO)
Small convolution code	0.25µm	20mW 3.3V	160 Mbits	0.125nJ/b
$(16,11)^2$ Product Code	0.18µm	7mW 1.8V	100 Mbits	0.7nJ/b (core+ IO)
(8,4) Low Voltage	0.18µm	0.036mW 1.8V	4.4 Mbits	0.008nJ/b
(8,4) trellis (low voltage)	0.18µm	150µW 0.8V	3.7 Mbits	0.042nJ/b
(8,4) trellis (low voltage)	0.18µm	2.4µW 0.5V	69 kbits	0.034nJ/b

The tables clearly show that the analog iterative decoders consume low energy per bit and can replace digital decoders. Low voltage design techniques [WINS06] lead to low power consumption while still maintaining high throughput.

The achievable coding gain can range from 3dB to 6dB for iterative decoders. The BER can be as low as 10^{-5}. Inclusion of these decoders can be of significant benefit to the RF subsytem of the µNode project. By increasing the coding gain by 3dB or decreasing the BER, the transmit power required in the power amplfier can be reduced to almost half. Keeping the power in power amplifier constant, we can extend the range of communication or relax the constraints on antenna gain. This can all be done because the excellent error correcting nature of these iterative codes allows for more errors in the system. Also using error control codes can help in decreasing the packet error rate thereby reducing retransmissions and hence saving more energy over time.

A DECODER EXAMPLE

To implement these iterative decoding algorithm a generic algorithm called the sum-and-product algorithm (SPA) is used (Loeliger, et al. 2001). The SPA is used to implement a general framwork for probabilty propagation. The SPA algorithm implements constraint graphs which express the logical relation between bits of the codeword. It computes global probabilites using local constraints and is given by:

$$P(z = j) = \eta \sum P(x = k)P(y = l)$$

Decoding algorithms such as trellis decoding using the BCJR algorithm [BAHL74], the turbo products can be shown to be a subset of the SPA algorithm. The SPA algorithm is very neatly implemented by taking advantage of the exponential relationship between gate bias and current of a CMOS transistor operating in the subthreshold region [LOEL01].

Popular circuits such as the Gilber Vector Multiplier operating in the subthreshold region are used to multiply probabilties and the necessary sums are computed simply by shorting wires together. These circuits implement parity check nodes, equality nodes and other types of operations that are required for implementing these decoders.

Input to these decoders can either be current or voltage with probabilites expressed in the Log Likelihood Ratio domain. The Gilbert Vector Multiplier implments for two current vectors x and y:

$$Z = \frac{yx^T}{\sum_k x_k}$$

where the current vectors x and y can be designed to be proportion to probabilty masses.

To explore the implementation of this decoding technique in 65nm technology a simple [5,2,3] linear block code was simulated. This example is outlined in detail in [LUST00]. Iterative codes generally consist of such suboptimal but more complex codes connected to each other with some constraints. Hence it is suitable to try and simulate a simple component decoder of these iterative codes.

The encoder takes two bits at a time and converts them in 5 bits with a Hamming distance of 3 between each code. A trellis structure for this linear block code is devised and then trellis decoding is implemented using the BCJR algorithm[BAHL74]. The output of the decoder are currents proportional to probabilities masses of each of the two bits being a 1 or 0. A hard decision then can be taken on these current outputs to decide with more confidence as to what bit was sent.

The power supply was kept at 1.2 V and the bias current at 100nA. A total of just 126 CMOS transistors were used to implement the decoder. The power consumption was just 1.2 µW for 100nA bias current and 0.7 µW for 50nA bias current. The lower power consumption for implementing this decoder is encouraging for implementing a more complex iterative decoder. One thing to note here is that, even if the results of analog demodulation and error decoding are not precise, the error correcting nature of the iterative codes takes care of moderate numbers of errors introduced into the received data.

ISSUES AT 65NM FOR ANALOG ITERATIVE DECODERS

Short channel effects such as drain induced barrier lowering, channel length modulation and mismatch lead to variation in subthreshold slope and threshold voltage. There is almost a linear relationship between DIBL and threshold voltage variation. However there is an exponential relationship between threshold voltage variation and current in the subthreshold region which severely degrades the probability calculation of these decoders. The simple solution is to use large length transistiors, but this can lead to slow speeds due to increased parasitic capacitances. Apart from large lengths, better mirroring techniques need to be deployed for better matching and lowering drain bias dependence. [ZARG08] shows simulated and measured data related to 65nm technology. It shows that mismatch and DIBL can severly affect the perfomance of these analog decoders, and special care needs to be taken while implementing these circuits.

With reduced power supply voltage, headroom also becomes an issue with large stacked stages of decoders. Folding can be used to reduce the stack size, but it increases power consumption and decreases speed. Low voltage techniques and mismatch analysis techniques listed in [WINS05] are crucial for implementing analog iterative decoders in 65nm technology.

DCD AND DSSS IN ANALOG

Implementing DCD and DSSS correlation in analog can also be an exciting avenue to explore to reduce power consumption. Morever as stated earlier, the output of these blocks would be discretely sampled analog values which can used as soft inputs to an iterative analog decoder. Currently to demodulate DCD DSSS encoded data, first a decoding of DCD is done and then "de-spreading" of DSSS modulated data is done using digital cross-correlators. This is the most power hungry process in the whole digital baseband system.

Decoding DCD involves taking products of samples from the demodulated inphase and quadrature phase components and then subtracting. This can be easily achieved by using the sum-product-algorithm which was used above to implement the analog decoders.

There has been previous work on implementing analog correlators using bank of capacitors with promising results(Onodera & Gray, 1998). In the future course of research this idea will be further explored for implementation and fabrication. The timing recovery and synchronization can still be performed in digital and can be used to control the analog correlator. This would require a use of analog to digital converter but only running at the data rate and not at the chip rate. However, research shows that it may also be possible to implement synchronization schemes in analog [DAUW04].

FUTURE WORK

1. Implement a more complex iterative decoder and evaluate perfomance in terms of power, area, SNR, BER, coding gain, mismatch etc.
2. Implement a DCD decoding algorithm
3. Implement an analog correlator for performing the "de-spreading" of the DSSS-modulated data and integrating it with digital synchronization and timing recovery circuits.

REFERENCES

[BAHL74] Bahl, L. R., J. Cocke, F. Jelnek, and J. Raviv. "Optimal Decoding of Linear Codes for Minimizing Symbol Error Rate." IEEE Transaction of Information Theory, March 1974: 284-287.

[BERR93] Berrou, C., A. Glavieux, and P. Thitimajshima. "Near Shannon limit error correcting coding and decoding: Turbo codes." International Communications Conference, May 1993: 1064–1070.

[CAVA97] A. Cavallini, F. Giannetti, M. Luise, and H. T. Nguyen, "Chip-level differential encoding/detection of spread-spectrum signals for CDMA radio transmission over fading channels," IEEE Transactions on Communications, Vol. 45, No. 4, pp. 456-463, April 1997.

[COLA02] G. Colavolpe and B. Raheli, “Improved differential detection of chip-level differentially encoded direct sequence spread-spectrum signals,” IEEE Transactions on Wireless Communications, Vol. 1, No. 1, pp. 125-133, Jan. 2002.

[COUC93] L. W. Couch II, Digital and Analog Communication Systems, 4th Ed., New York: Macmillan Publishing Company. 1993.

[DAUW04] Dauwels, Justin, Matthias Frey, Tobias Koch, Hans-Andrea Loeliger, Patrick Merkli, and Benjamin Vigoda. "Synchronization of a Pseudo-Noise Signal using an Analog Circuit." Technical Report No 200401, Signal and Information Processing Laboratory, Swiss Federal Institute of Technology, Zurich, 2004.

[ELIA54] Elias, P. "Error-free coding." IRE Trans. on Information Theory IT-4 (September 1954): 29-37.

[GALL63] Gallager, R. G.; Low-Density Parity Check Codes. Cambridge, MA: MIT Press, 1963.

[GORD04] P. Gorday, Q. Shi, and F. Martin, “Performance of Chip-Level Differential Detection with Phase Noise,” IEEE Wireless Communications and Networking Conference (WCNC), March 21-25, 2004.

[HAGE96] Hagenauer, Joachim, Elke Offer, and Lutz Papke. "Iterative Decoding of Binary Block and Convolutional Codes." IEEE Transaction on Information Theory 42, no. 2 (March 1996): 429-445.

[HOGE81] Hogenauer, IEEE Trans. on Acoustics, Speech, and Signal Processing, April 1981

[LOEL01] Loeliger, H. A., F. Lustenberger, M. Helfenstein, and F. Tarkoy. "Proabibilty propogation and decoding in analog VLSI." IEEE Transaction on Information Theory 47, no. 2 (February 2001): 837-843.

[LUST00] Lustenberger, F. "On the Design of Analog VLSI Iterative Decoders." PhD Thesis, 2000.

[MEYS08] Zargham Meysam, G. V. (2008, July 12). Scaling Analog Decoders to 65-nm CMOS Technologies. Retrieved from The 2008 International Analog Decoding Workshop: http://www.neng.usu.edu/ece/faculty/winstead/adw08/Gaudet_65nm_Decoders_ADW08.ppt

[ONOD98] Onodera, Keith K., and Paul R. Gray. "A 75 mW 128-MHz DS-CDMA Baseband Demodulator for High-Speed Wireless Applications." IEEE Journal of Solid-State Circuits 33, no. 5 (May 1998).

[SCHL07] Schlegel, Christian; Analog Processing of Digital Data: Promise and Challenges. July 11-13, 2007. www.ece.ualberta.ca/~schlegel/Pub/CMOSwkshop07.pdf.

[SHAN48] Shannon, Claude. "A mathematical theory of communication." Bell System Technical Journal, July 1948.

[SHI02] Q. Shi, R. J. O’Dea, and F. Martin, “A new chip-level differential detection system for DS-CDMA,” IEEE International Conference on Communications (ICC), Vol. 1, pp. 544-547, April 2002.

[VUCE00] Vucetic, Branka, and Jinhong Yuan. Turbo Codes. Kluwer Academic Publishers, 2000.

[WINS05] Winstead, Chris. "Analog Iterative Error Control Decoders." Phd Thesis, 2005.

[WINS06] Winstead, Chris, Nguyen Nhan, Vincent C. Gaudet, and Schlegel Christian. "Low-voltage CMOS Circuits for Analog Iterative Decoders." IEEE Transaction on Circuits and System-I: Regular Papers 53, no. 4 (April 2006): 829-841.

[ZARG08] Zargham Meysam, Gaudet Vincent, Schlegel Vincent. "Scaling Analog Decoders to 65-nm CMOS Technologies." The 2008 International Analog Decoding Workshop. July 12, 2008. http://www.neng.usu.edu/ece/faculty/winstead/adw08/Gaudet_65nm_Decoders_ADW08.ppt.