

# Energy Efficient Modulation and MAC for Asymmetric RF Microsensor Systems

Andrew Y. Wang, SeongHwan Cho, Charles G. Sodini, Anantha P. Chandrakasan

Department of Electrical Engineering and Computer Science  
Massachusetts Institute of Technology, Cambridge, MA

{andywang,chosta,sodini,anantha}@mit.edu

## ABSTRACT

Wireless microsensor systems are used in a variety of civil and military applications. Such microsensors are required to operate for years from a small energy source. To minimize the energy dissipation of the sensor node, RF front-end circuitry must be designed based on system level optimization of the entire network. This paper presents several energy minimization techniques derived from the unique properties of a practical short range asymmetric microsensor system. These include energy efficient modulation schemes, appropriate multiple access protocols, and a fast turn-on transmitter architecture.

## 1. INTRODUCTION

The design of wireless microsensor systems have gained increasing importance for a variety of civil and military applications. With the objective of providing short-range connectivity with significant fault tolerances, these systems find usages in diverse areas such as environmental monitoring, industrial process automation, and field surveillance.

Wireless microsensor systems are quite distinctive from conventional data and voice applications. First, a microsensor cell covers a small area (meters vs. km). Second, sensors have very low data rate, which is typically on the order of a few hundred bits per second. Third, the data link is highly asymmetric since most of the data flow from the sensor nodes to the base station. Fourth, delay is usually a stringent requirement in real time applications. Finally, reliability (low message error rate), size, and ease of deployment are all important design considerations. Table 1 shows a detailed specification for a sensor system that is used in a factory machine monitoring environment.

The main design objective is to maximize the battery life time of the sensor nodes while ensuring reliable operation. For many applications, the sensors need to operate for 5-10 years without battery replacement. To achieve this goal, the microsensor system has to be designed in a highly integrated fashion and optimized across all levels of system abstraction.

cell density	200 - 300 in 5mx5m area 2000 - 3000 in 100mx100m area
range of link	< 10m
message rate (msg = 2bytes)	average: 20 msgs/sec maximum: 100 msgs/sec minimum: 2 msgs/sec
error rate and latency	$10^{-6}$ after 5ms $10^{-9}$ after 10ms $10^{-12}$ after 15ms
battery life	5-10 years
size	one AA size battery

Table 1: Wireless microsensor system specification for an example machine monitoring application

This also means that all the characteristics particular to the microsensor system must be exploited.

Over the past years, there has been active research in microsensor networks at both the system and the circuit levels. At the system level, a number of interesting results have been published on routing and flow control algorithms that increase the Quality of Service (QoS) of ad-hoc networks [3, 11, 4]. At the circuit level, the main focus has been on the design of low-power high sensitivity RF front-end circuitry. Various techniques have been applied to increase the performance of individual front-end components [2, 1, 6].

There has been little work done in the design of modulation techniques and multi-access (MAC) protocols, optimized for short range asymmetric wireless microsensor networks, from the standpoint of circuit power consumption. In this paper we will show that RF front-end circuitry must be designed based on system-level optimization of the entire network in order to achieve minimal energy dissipation. First we will demonstrate the importance of reducing transmitter start-up time, and then we will describe energy efficient modulation schemes and multi-access protocols suitable for wireless sensor systems.

## 2. PROBLEM OF START-UP ENERGY

The communication module of a wireless sensor, as shown in Figure 1, must be designed for low duty cycle activity. For short range transmission at GHz carrier frequency, the power is dominated by the radio electronics (frequency synthesizer, mixers, etc.) [10], which is on the order of 10-100mW [8, 1]. The output transmit power, on the other hand, is only about 0dBm (1mW) for bit error rate (BER)

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

ISLPED'01, August 6-7, 2001, Huntington Beach, California, USA.

Copyright 2001 ACM 1-58113-371-5/01/0008 ...\$5.00.

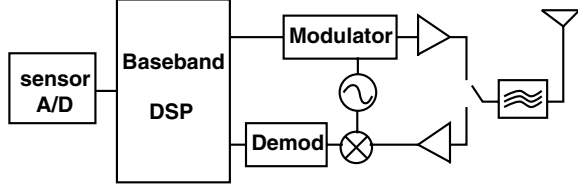


Figure 1: Generic Microsensor Transceiver

as low as  $10^{-5}$  at 1Mbps. Moreover, the power consumption of the transceiver does not vary much with the data rate to the first order [10]. Therefore, it makes sense to send the packets in burst at high data rate to minimize the transmission time and shut off the transmitter during the idle period. Unfortunately, transmitters require a significant overhead in terms of time and energy dissipation to go from the sleep state to the active state. Typical start-up time is on the order of  $100\mu s$  or more, while the transmit on-time is less than that. This means that the transient energy during the start-up can be higher than the energy required by the electronics during the actual transmission.

The effect of start-up transient is shown in Figure 2, where energy consumption per packet is plotted against start-up time [3]. The dotted line indicates the energy consumption if there was no start-up transient. The solid line shows that as the start-up time increases, the energy consumption is dominated by the start-up transient and not by the transmit on-time. Hence it is essential to minimize the start-up time in order to reduce power consumption. We will discuss a fast start-up transmitter architecture in Section 5.

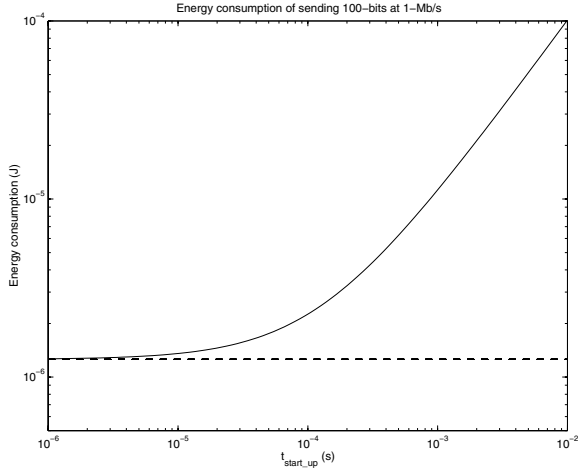


Figure 2: Effect of start-up transient

### 3. MODULATION TECHNIQUES

In this section we analyze how to reduce the energy dissipation during the on-time of the transmitter. A generic In-phase/Quadrature (I/Q) transmitter architecture is shown in Figure 3. This architecture is necessary for PSK and QAM modulators. However, for FSK, it is possible to employ a direct modulation approach such as an open loop VCO modulator or a closed-loop architecture that does not

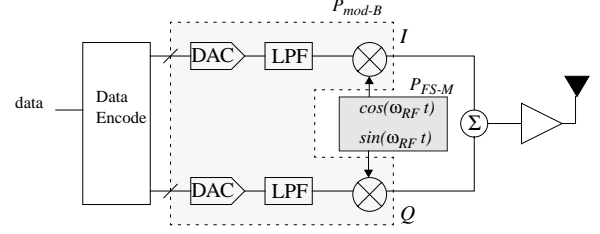


Figure 3: Generic I/Q transmitter

require DACs, mixers, and quadrature VCOs as described in Section 5.

A simple model for the on-time energy dissipation is:

$$E_{on} = (P_{TX} + P_{RF}) \cdot t_{on} \quad (1)$$

where  $P_{TX}$  is the power dissipation of the transmitter electronics, and  $P_{RF}$  is the RF output power.  $P_{RF}$  is typically small as compared to  $P_{TX}$ . We will discuss two techniques to achieve energy savings: reducing  $t_{on}$  through M-ary modulation and trading off  $P_{RF}$  with  $P_{TX}$ .

#### 3.1 Binary versus Multi-level Modulation

In M-ary modulation, each transmitted symbol comes from a set of  $M$  (rather than 2 as in binary) distinct waveforms. This means that  $\log_2 M$  bits are sent per symbol. Examples are M-ary Phase Shift Keying (M-PSK), M-ary Quadrature Amplitude Modulation (M-QAM), and M-ary Frequency Shift Keying (M-FSK). Figure 4 shows a plot of required transmit power versus bandwidth efficiency, which is defined as bit rate per required bandwidth. This result comes from a link-budget analysis with a carrier frequency of 5.8GHz, a BER of  $10^{-5}$ , and a data rate of 1MSymbols/s. Rayleigh fading is assumed, which is a fairly good model for indoor channels [9, 5]. Large scale fading is assumed to have a path loss exponent of  $2 \sim 3$ , which conforms to an indoor factory environment with a large number of metal equipment.

Note the distinction between M-PSK/M-QAM, which sacrifices transmit power to obtain higher bandwidth efficiency (typically to accommodate more users for a fixed amount of bandwidth), and M-FSK, which sacrifices bandwidth for a reduction in transmit power (in terms of transmit power per bit). Thus power efficient modulation techniques such as FSK should be considered for applications where power minimization is the main objective.

If the symbol rate is kept constant at 1MSymbols/s, then M-ary modulation reduces transmit on-time by a factor of  $\log_2 M$  because M-ary modulation produces  $\log_2 M$  bits per symbol. However, this comes at a cost of increased  $P_{TX}$  as well as  $P_{RF}$ . In order to compare the performance of binary versus M-ary transmitters, the following models are used:

$$E_{2PSK} = P_{FS} \cdot t_{start} + (P_{MOD} + P_{FS} + P_{RF}) \cdot t_{on} \quad (2)$$

$$E_{MARY} = \beta P_{FS} \cdot t_{start} + (\alpha P_{MOD} + \beta P_{FS} + \gamma P_{RF}) \cdot t_{on}/r \quad (3)$$

where in the first equation,  $E_{2PSK}$  is the energy consumption per transmitted packet for a binary PSK transmitter,  $P_{FS}$  is the power consumption of the frequency synthesizer, which is dominant,  $P_{MOD}$  is the power consumption of all

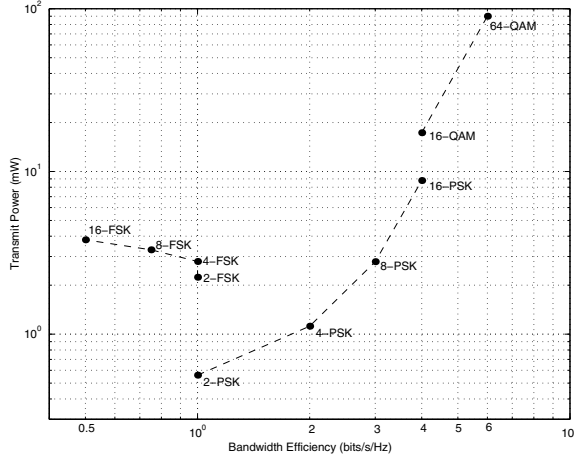


Figure 4: Transmit power versus bandwidth efficiency in fading channel

the other transmitter circuitry (i.e.,  $P_{MOD} = P_{TX} - P_{FS}$ ), and  $P_{RF}$  is the RF transmit power. This equation is appropriate for a short-range microsensor system but may not be suitable for other systems such as cellular where the power amplifier is the dominant source of power consumption. Equation 3 describes the energy dissipation for an M-ary transmitter.  $E_{MARY}$  is written in terms of the variables used in  $E_{2PSK}$  for the purpose of easy comparison. The Greek alphabets represent the extra overhead energy, or cost, required for M-ary modulation systems:  $\alpha$  is the overhead in DAC, LPF, mixers, etc.,  $\beta$  is the overhead in the frequency synthesizer, and  $\gamma$  is the overhead RF power which can be computed from Figure 4.  $r$  equals to  $\log_2 M$ .

We estimated that  $P_{FS} = 10\text{mW}$ ,  $P_{MOD} = 2\text{mW}$ ,  $\alpha = 1 \sim 3$ ,  $\beta = 1.75$ , and  $t_{on} = 100\mu\text{s}$ . Although  $P_{FS}$  and  $P_{MOD}$  values above are aggressive as compared to what are commercially available today, they can be achieved through advanced technology scaling and novel techniques in architecture.

M-ary modulation is more energy efficient than binary modulation when  $E_{MARY} < E_{2PSK}$ . Figure 5 shows the relative energy savings. Energy savings for M-ary modulation decreases as  $t_{start}$  increases. This is because when  $t_{start}$  is long, the start-up energy of the frequency synthesizer dominates, so the energy savings gained through the reduction of  $t_{on}$  are negligible. As  $t_{start}$  becomes shorter, the on-time energy dissipation becomes the dominant term, so reducing  $t_{on}$  through M-ary modulation achieves significant energy savings. Therefore, reducing  $t_{start}$  not only decreases the start-up energy  $E_{start}$  but also helps M-ary modulation to reduce the on-time energy  $E_{on}$ . To summarize, M-ary modulation achieves the greatest energy savings when the ratio  $t_{on}/t_{start}$  is large and  $P_{RF}$  is small (relative to  $P_{FS}$ ). Since  $t_{on}$  is usually determined by the data rate, it is important to minimize  $t_{start}$  and  $P_{RF}$  to make M-ary modulation even more energy efficient.

The second important observation is that M-FSK becomes more efficient than M-PSK/M-QAM for  $M > 8$ . M-FSK is not as energy-efficient at small M because noncoherent detection requires 6dB more RF power to achieve the same BER performance. For large M, the symbol SNR required for M-PSK/M-QAM grows very fast, which offsets the en-

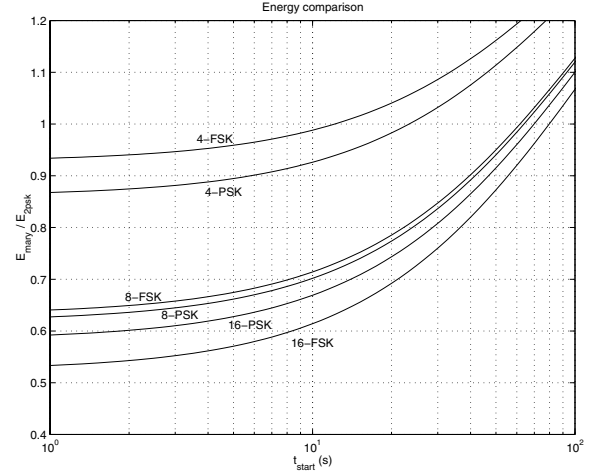


Figure 5: Energy savings as a function of start-up time

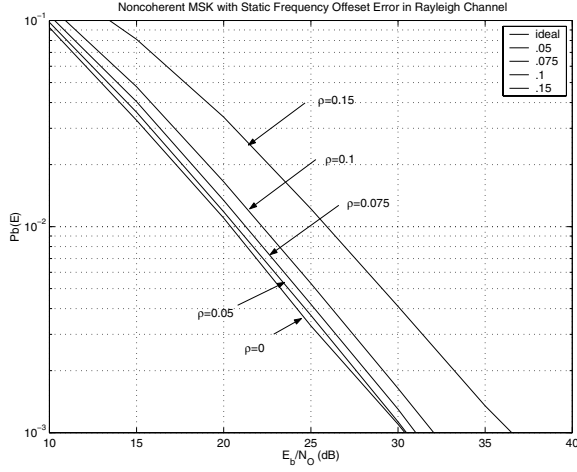
ergy savings gained through reduction of  $t_{on}$ . The symbol SNR required for M-FSK grows slowly, thereby making it very energy-efficient at large M. This makes M-FSK attractive since M-FSK already has the advantage of not requiring carrier synchronization. The sacrifice, however, is bandwidth. For instance, 8-FSK uses 4 times as much bandwidth as M-PSK. This problem may be circumvented by careful planning of the spectrum. In the unlicensed band in the GHz regime, large bandwidth is available to make M-FSK a realistic option.

### 3.2 Reducing Transmitter Complexity

The transmitter electronics power can be lowered by reducing the performance requirements of critical transmitter components – for example, the phase noise requirement of the VCO and the frequency offset error of the frequency synthesizer.

The phase noise of the VCO and the frequency offset error of the frequency synthesizer create two major concerns. The immediate impact is degradation of performance in bit error rate. A phase tracking error occurs due to phase noise, frequency error, and non-ideal frequency response of the phase-locked loop, in addition to I/Q mismatch created by quantization and gain errors. The second concern, which may be more serious, is that large phase error caused by phase noise and frequency error can potentially cause the carrier tracking loop to lose lock in coherent systems. This problem is exacerbated in a fading channel where carrier synchronization is usually a difficult task. It has been shown that the modulation index error of Minimum Shift Keying (MSK) has to be kept below 5% to achieve a reasonable RMS phase tracking error if an aggressive carrier tracking loop bandwidth of approximately 1% of the symbol rate is used [7]. This is a very stringent restriction. For example, the Digital Enhanced Cordless Telecommunications (DECT) standard specifies a 10% accuracy in modulation index, which is not adequate for use with coherent detection.

In light of the above observation, noncoherent detection provides an attractive alternative since it does not require carrier phase tracking. Figure 6 shows the effect of frequency error on the bit error rate of noncoherent MSK in Rayleigh



**Figure 6: MSK BER degradation due to frequency offset error in Rayleigh fading channel**

channel.  $\rho$  is the normalized frequency error and is defined as  $\rho = f_e T$ , where  $f_e$  is the actual frequency error, and  $T$  is the symbol period.

As shown in Figure 6, performance degradation is not severe even for moderately large phase and frequency errors. It takes about 2dB of  $E_b/N_o$  to compensate for a frequency error of  $\rho = 0.1$ , which corresponds to a 20% modulation index error.

This suggests that it is possible to reduce the transmitter energy consumption by increasing the RF output power to compensate for more relaxed phase noise and frequency error requirements. Specifically, Equation 3 is modified in the following way,

$$E_{MARY} = (1 - \delta)\beta P_{FS} \cdot t_{start} + [\alpha P_{MOD} + (1 - \delta)\beta P_{FS} + (1 + \kappa)\gamma P_{RF}] \cdot t_{on}/r \quad (4)$$

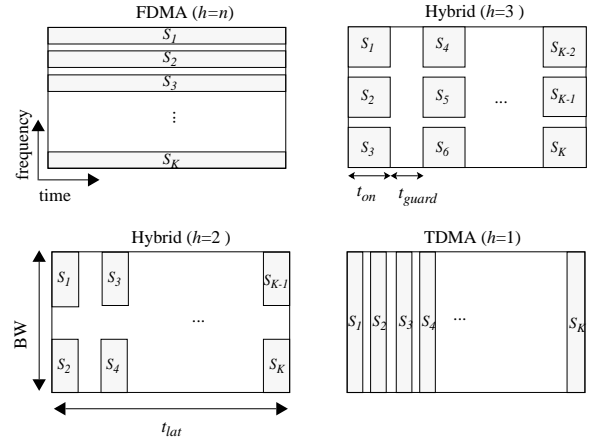
where  $\delta$  is the reduction in the frequency synthesizer power due to relaxed phase noise and frequency error, and  $\kappa$  represents the increase in RF output power that compensates the BER loss. The overall energy consumption is lowered if

$$\delta \geq \frac{\kappa\gamma P_{RF} t_{on}/r}{\beta P_{FS} (t_{start} + t_{on}/r)} \quad (5)$$

Energy savings depend on the values of  $\delta$  and  $\kappa$  and are on the order of a few tens of percent.

#### 4. MULTI-ACCESS PROTOCOLS

In this section we derive a low power MAC protocol based on the the model that is extracted from the physical layer electronics. The average power consumption of a sensor radio is given by the following equation, where  $N_{tx/rx}$  is the average number of times per second that the sensor transmitter/receiver is used, which is specified by the application scenario and communication protocol,  $P_{TX/RX}$  is the power consumption of the transceiver,  $t_{on-tx/rx}$  is the transmit/receive on-time (actual data transmission/reception time),  $t_{start}$  is the start-up time of the transceiver, and  $P_{RF}$  is the output transmit power which drives the antenna.



**Figure 7: Multiple access methods**

$$P_{AVG} = N_{tx} \{ P_{TX} (t_{on-tx} + t_{start}) + P_{RF} t_{on-tx} \} + N_{rx} P_{RX} (t_{on-rx} + t_{start}) \quad (6)$$

If the bandwidth required by the sensor network is less than the available bandwidth, then FDMA offers the simplest MAC solution. However, when there are a large number of sensors, the available bandwidth may not be enough for each sensor to transmit at a high burst rate using FDMA. In this case, we need to find an alternative solution that is more energy-efficient. The following describes how such a solution can be determined.

In TDMA,  $t_{on}$  is minimized since the bandwidth is at maximum, allowing the highest data rate. For FDMA, the available bandwidth is at minimum, resulting in the longest on-time. A hybrid scheme of TDM-FDM is also possible, where both time and frequency are divided into transmission slots. This is illustrated in Figure 7 where shaded area indicates a valid transmit slot for sensor  $S_i$ . In cases where time division is employed, we should note that a downlink from the basestation to the sensors is required in order to maintain time synchronization among the sensors. Due to the finite error among each sensor's reference clock, the basestation must send out *sync* signals as to avoid any collision among the transmitted packets. Hence the sensor receiver must be turned on every so often to receive these *sync* packets. The number of sync packet receptions ( $N_{rx}$ ) depends on the guard time ( $t_{guard}$ ) which is the minimum time difference between two time slots in the same frequency band, as shown in Figure 7. If two slots in the same frequency band are separated by  $t_{guard}$ , it will take  $t_{guard}/\varepsilon$  time for these two packets to collide, where  $\varepsilon$  is the difference between the two reference clocks. Hence the sensor must be resynchronized at least  $\varepsilon/t_{guard}$  number of times every second. This is described in Eq. 7, where  $BW$  is the total available bandwidth,  $Data$  is the size of the transmit packet in bits,  $t_{lat}$  is the latency requirement of the sensor data,  $h$  is the number of channels in the given  $BW$ ,  $t_{avail}$  is the time difference between start of two packets (Figure 7), and  $K$  is the number of sensors. It is also assumed that the data rate is equal to the occupying signal bandwidth and hence  $t_{on} = Data/(BW/h)$ .

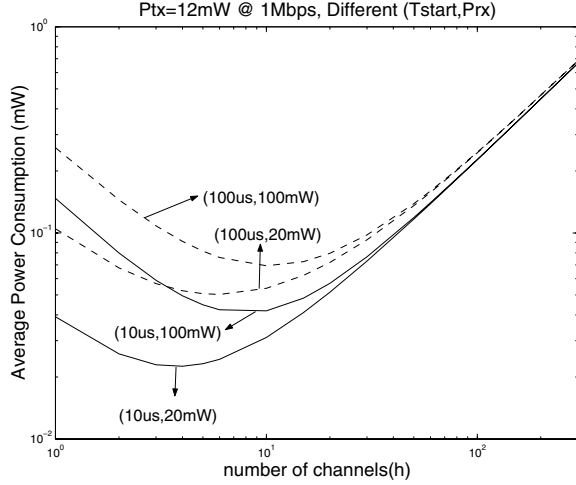


Figure 8: Energy with different  $(t_{start}, P_{RX})$

$$N_{rx} = \frac{\epsilon}{t_{guard}} = \frac{\epsilon}{t_{avail} - t_{on}} = \frac{\epsilon}{(\frac{t_{lat}}{K} - \frac{Data}{BW})h} \quad (7)$$

From the above equation, we see that as the number of channels decreases, guard time becomes larger and receiver activity is reduced. It is also apparent that the advantage of pure FDMA is that it does not need a receiver (i.e.  $t_{guard} \rightarrow \infty, N_{rx} = 0$ ).

A simulation is performed for a scenario where a sensor on average sends twenty 100-bit packets/sec ( $N_{tx} = 20/sec, Data = 100bits$ ) with 5ms latency requirement ( $t_{lat} = 5ms$ ). The bandwidth available to the cell is 10MHz ( $BW = 10MHz$ ), and the number of sensors is 300. The resulting average power consumption is plotted in Figure 8, where average power consumption is plotted versus the number of channels (i.e.,  $h = 1$ : TDMA,  $h = 300$ : FDMA). The graph shows power consumption for different  $P_{RX}/P_{TX}$  and  $t_{start}$ . It can be seen that the optimal number of channels that results in the lowest energy increases for higher receiver power, as to reduce the number of receptions. The reason why TDMA with minimum on-time does not achieve the lowest power is because of the receiver power consumption from network synchronization. As the number of channels increase, guard time becomes smaller and the receiver power starts to become a significant portion of overall power consumption.

## 5. CONSIDERATIONS ON FAST START-UP LOW POWER FREQUENCY SYNTHESIZER

It has been seen from the previous sections that a fast start-up low power transmitter is necessary for an energy efficient transmission. An attractive transmitter architecture that meets these goals is based on a fractional-N (FN) frequency synthesizer with  $\Sigma\Delta$  modulator. There are several advantages of this architecture over other transmitter architectures. First, FN synthesizer allows higher reference frequency than integer-N architectures since channel spacing is not limited to integer multiples of the reference frequency. Consequently, a higher loop bandwidth is available

and hence faster start-up time can be achieved. Second, FSK modulation is readily available by exploiting the noise shaping properties of the  $\Sigma\Delta$  modulator. This leads to lower power consumption in the overall transmitter since mixers, DACs, and quadrature VCOs are not necessary. The achievable start-up time and data rate is governed by the loop bandwidth of the synthesizer. Basically, higher loop bandwidth allows both faster start-up time and higher data rate. For example, loop bandwidth of 700kHz allows 1Mbps FSK modulation with start-up time of less than  $10\mu s$  [12].

From the standpoint of power consumption and output noise, increasing the loop bandwidth of the synthesizer has two adverse effects which can be exploited to reduce the synthesizer's overall power consumption.

The output noise of a silicon integrated synthesizer is usually dominated by the phase noise of the VCO and the quantization noise from the  $\Sigma\Delta$  modulator. The noise from each source goes through different loop characteristics of the PLL and shows up at the output of the synthesizer as shown in Fig. 9. Specifically, the phase noise from the VCO goes through high pass characteristic of the loop while the quantization noise from the  $\Sigma\Delta$  goes through low pass characteristic of the loop. This can be described in Equation 8 where  $\Phi_{VCO}$  is the phase noise of the VCO,  $\Phi_{\Sigma\Delta}$  is the quantization noise from the  $\Sigma\Delta$ ,  $\omega_c$  is the loop bandwidth of the synthesizer, and  $f_1(\omega_c)$ ,  $f_2(\omega_c)$  are the loop transfer functions seen from the VCO and the  $\Sigma\Delta$ , respectively.

$$\Phi_{total} \simeq f_1(\omega_c)\Phi_{VCO} + f_2(\omega_c)\Phi_{\Sigma\Delta} \quad (8)$$

In detail,  $\Phi_{VCO}$  and  $\Phi_{\Sigma\Delta}$  can be described by the following equations, where in Equation 9,  $P_{VCO}$  is the power consumption of the VCO,  $Q$  is the quality factor of the tank,  $2FkT$  is the thermal noise constant,  $f_3$  is the flicker noise corner of the device and  $f$  is the operating frequency. In Equation 10,  $f_{ref}$  is the operating frequency of the  $\Sigma\Delta$  and  $m$  is the order of the  $\Sigma\Delta$ . It should also be noted that power consumption of the  $\Sigma\Delta$ ,  $P_{\Sigma\Delta}$ , is proportional to its order  $m$  and  $f_{ref}$ .

$$\Phi_{VCO} = \frac{2FkT}{P_{VCO}} \left\{ 1 + \frac{f_o}{2Qf} \right\} \left\{ 1 + \frac{f_3}{f} \right\} \quad (9)$$

$$\Phi_{\Sigma\Delta} = \frac{(2\pi)^2}{12f_{ref}} \left\{ \sin\left(\pi \frac{f}{f_{ref}}\right) \right\}^{2(m-1)} \quad (10)$$

$$P_{\Sigma\Delta} \propto mf_{ref} \quad (11)$$

Since each noise source goes through different loop characteristics, power consumption can be reduced by adjusting the loop bandwidth so that the noise requirement of that component is less stringent. For example, we can tolerate a higher phase noise VCO and yet still achieve the same output noise by increasing  $\omega_c$ . By allowing higher VCO phase noise, we can lower the power consumption of the VCO as seen in Equation 9. From the perspective of the quantization noise from the  $\Sigma\Delta$ , increase in  $\omega_c$  has an opposing effect on the output noise. Since  $P_{\Sigma\Delta}$  is high pass filtered through the PLL, higher  $\omega_c$  increases the output noise from the  $\Sigma\Delta$ . If the overall output noise is kept to be the same, power of the  $\Sigma\Delta$  must be increased accordingly. Therefore we see a trade-off between the VCO and the  $\Sigma\Delta$  power consumption as the loop bandwidth varies. A large loop bandwidth

