Wireless Handset RF Front-End Optimization

October 2014

TABLE OF CONTENTS

1. INTRODUCTION

1.1 EXECUTIVE SUMMARY

The goal of this paper is to educate wireless stakeholders on the issues, challenges, complexity and opportunities in regard to wireless handset Radio Frequency (RF) Front-End optimization without "overhyping" or "overpromising" any future solutions.

The exponential growth of mobile data has driven up the need of more wireless spectrum. As diverse and fragmented spectrum bands that span from low to high need to be supported on a device, it presents a dramatic challenge to handset RF front-end design and architecture. To further this challenge, market forces are also driving handsets to meet the following complex requirements: multi-band (for domestic & international roaming) and multi-mode (2G/3G/4G, Assisted Global Positioning System (AGPS), WiFi/Bluetooth (BT), Near Field Communications (NFC)), Multi-Input Multi-Output (MIMO) and carrier aggregation, small form factor (thinner ID, larger display, etc.) while balancing cost competitiveness, ever better performance, longer battery life and regulatory requirements (i.e., Specific Absorption Rate (SAR)). As such, RF Front-End must be optimized for the benefit of stakeholders and end customers. An optimized RF Front-End design and architecture can be very helpful to meet all these challenges and facilitate adoption of 3rd Generation Partnership Project (3GPP) mobile broadband technologies in all appropriate frequency bands in North America and other parts of the world.

This white paper discusses the importance of an optimized RF Front-End design and architecture to support new and fragmented spectrum bands proposing solutions without sacrificing performance of the device. The white paper also identifies opportunities with current RF Front-End design and architecture and discusses potential solutions. Different Hardware/Software (HW/SW) solutions have advantages and/or limitations and ultimately spectrum harmonization remains an efficient way to optimize the RF Front-End. The paper also serves the purpose of educating key stakeholders such as press, analysts, carriers, vendors, device manufacturers and application developers. This paper illustrates, for the non-RF expert, the changes that are occurring in the RF Front-End as the Mobile Radio System Generation evolves to 4G and beyond. Details on the higher-level trends in the Mobile Radio System will not be discussed; rather, this paper will refer to existing documents to cover those details

There are four significant trends in wireless technology:

- Increasing data usage
- A desire for ubiquitous coverage
- Balanced Uplink (UL) and Downlink (DL) data rates
- User Equipment (UE) with bigger screens and batteries, but a smaller volume left over for the radio

These trends have driven changes in the design of the wireless network and in the design of the phone itself:

- There are more base stations in a given area (spatial densification)
- Each UE needs to have multiple antennas
- Each UE needs to operate simultaneously on multiple frequencies or channels, using more spectrum
- With higher order modulation, requiring a higher per-channel signal-to-noise ratio

But there are several limitations on what is achievable. Each receive or transmit chain needs at least some physical volume within the phone, and uses some significant amount of power. There is an upper bound on the number of paths the Radio Frequency Front-End (RFFE) can support. This upper bound is limited by size, cost and performance. The bound for technologies available today and in the foreseeable future has already been reached, causing User Equipment (UE) manufacturers to produce multiple Stock Keeping Units (SKU) for a given device, thus keeping popular devices limited to large markets. Research is underway to push these boundaries further, but the design and optimization of the RFFE is still a significant challenge.

1.2 DEFINITION OF RF FRONT-END

The RFFE in a UE is made up of a number of key components:

- The antenna(s) and antenna tuner(s)
- Band Select, Duplexers: filters, duplexers, diplexers, and switches used for frequency control
- Transmitters and RF Power Amplifiers (Tx/PA's)
- Receivers and Low-Noise Amplifiers (Rx/LNAs)

The Baseband and RF (mixers, down converter, etc.) section, a key component in the overall UE, is not part of the RFFE.

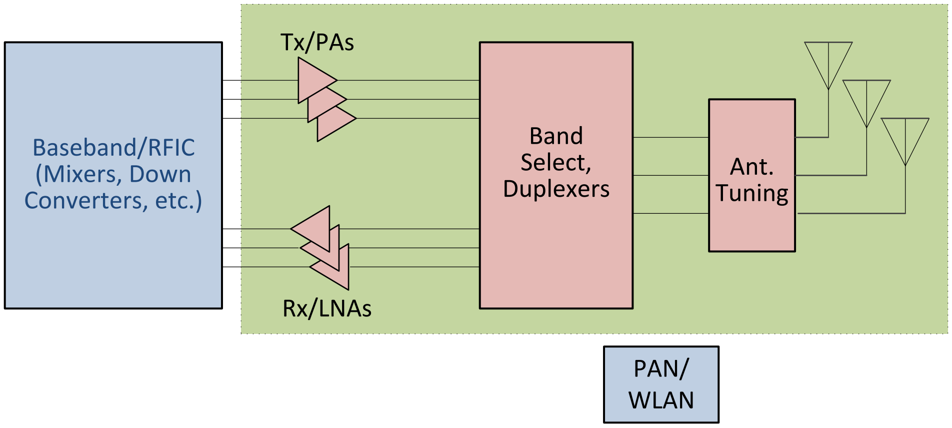

Figure 1. Simplified diagram of a UE, highlighting the RFFE portion.

1.3 DEFINITIONS OF TERMS / COMPONENTS / CONCEPTS

- **Base Transceiver Station (BTS):** Cellular base station, part of the Wireless Wide Area Network (WWAN) system
- **Diplexer:** A pair of interconnected filters (lowpass and highpass) used to separate groups of frequency bands, where the transition between the two filters does not have exceedingly stringent performance requirements.[1] There are also Triplexers and Quadplexers: these separate groups of frequency bands into 3 or 4, respectively. These often require higher performance than just a simple Diplexer.

---

[1] The Q of the filters is relatively low. See Wikipedia, "Q factor".

- **Duplexer:** A pair of high-quality filters interconnected in such a way as to separate the Transmit and Receive frequencies of a given frequency band[2]
- **Frequency Division Duplex (FDD):** Transmission and reception occur on different frequencies
- Low Noise Amplifier (LNA): The first RF amplifier in the receiver; it is usually connected directly to a duplexer or RF Filter
- **MIMO:** Using multiple antennas at both the transmitter and receiver
- **Power Amplifier (PA):** This document foc**uses on the RF Power** Amplifier.
- **PA Linearity:** A figure of merit for the amount of distortion (intermodulation or harmonic) created in the Power Amplifier of a transmitter. Higher Linearity is better.
- **Personal Area Network (PAN):** For example, Bluetooth
- **Receive Diversity:** Using multiple antennas at the receiver in order to pick up the best signal available at antenna output
- **Stock Keeping Unit (SKU):** a means to track items by a supplier. For example, a supplier might have a cellular phone of a given model number, with different RFFE's for sale in different regions. These two phones would have different SKU's.
- **Time Division Duplex (TDD):** At a given moment in time, a TDD device can either receive, or transmit, but not both.
- **Transmit Linearity:** A figure of merit for the amount of distortion (intermodulation, or harmonic) created in a transmitter. Higher Linearity is better.
- **UE:** For example, a mobile phone, tablet, or phablet.
- **Wide Area Network (WAN)**
- **Wireless Local Area Network (WLAN):** For example, WiFi
- **Wireless Wide Area Network (WWAN):** For example, the LTE network

2. BACKGROUND & MOTIVATION: DRIVERS FOR THE OPTIMIZATION OF RFFE

2.1 EVOLUTION OF MOBILE COMMUNICATION SYSTEMS AND DEVICES

Papers by 4G Americas, including ***4G Mobile Broadband Evolution: 3GPP Release 11 & Release 12 and Beyond***, and ***Meeting the 1000X Challenge: The Need for Spectrum, Technology and Policy Innovation***, present many of the motivations and changes that have occurred as wireless technology evolves. This paper will focus on the specific changes that affect the design of the RF Front-End (RFFE) in mobile handsets (User Equipment/UE). Although there are many different wireless standards and many different regulatory requirements that affect the details of a specific RFFE design, we will look at developments that affect the RFFE for all UEs.

In this section, we outline the relevant trends in wireless technology that have affected the RFFE in Mobile Handsets (UE). Next, in Section 2.2, we will look at how these trends have changed the design and optimization of the RFFE.

---

[2] The Q of the filters is quite high. The exact performance requirements are a function of the passband bandwidth, and the separation between receive and transmit frequencies: the closer Tx and Rx, the higher the Q required.

2.1.1. INCREASING DATA USAGE

Initially, the use of mobile handsets was exclusively for voice. But today, most mobile traffic is data. In fact, global mobile data traffic has doubled every year for the last few years. There are more and more subscribers, but also more and more data used by each subscriber. The shift from feature phones to smart phones and tablets with increasing screen sizes, increasing camera (still and video) resolutions, and improved user experience has driven an explosive growth in data usage. Users are moving towards being content creators as well as content consumers. This trend has far reaching implications for mobile networks and for the RFFE in mobile handsets; the data rates for the uplink must also increase a great deal, not just the downlink.[3]

The trend towards more data usage shows no signs of slowing. With newer applications (especially video) and more devices per user, the trend is expected to escalate.

As the peak data rate per user increases, the capacity of the network as a whole must also increase.

2.1.2. UBIQUITOUS COVERAGE

As data usage has increased, users expectation of being able to use applications no matter where they are has also increased. It is becoming unacceptable to have limited data rates in some areas compared to others (like at the edge-of-cell, or in coverage-holes).

In current systems, peak performance is only achievable within part of the coverage area of a given cell. Achieving close to full potential rates over the entire cell is even more of a challenge.

2.1.3. BIGGER SCREENS, BIGGER BATTERIES, SMALLER RFFE

Since the inception of the cellular phone, the major design trend has been to reduce its size. With the introduction of smartphone and phablet / tablet, this trend is changing; many new phones have physical dimensions much larger than previous generations and are largely driven by the screen size. However, this trend does not mean that there is more space in the phone. While the width and height of the phone have gotten larger, the thickness continues to decrease and more functionality has been added, like higher data rates, additional sensors and new technologies (for example, high resolution still and video cameras).

Additionally, users have demanded longer usage time from their devices, despite the increase in screen size and functionality. Battery density (capacity per unit of volume) typically only increases by 10 percent per year. The increased energy usage and usage time can only be accommodated with larger and larger batteries.

The result of these changes is that, even though more radio functionality is required, the volume allocated to the RFFE is staying the same or even shrinking.

---

[3] See Mobile Experts, 2012.

2.2 EVOLUTION OF RADIO FEATURES

2.2.1. PERFORMANCE EQUATION

The trends outlined above mean a need for greatly increased wireless system capacity, despite shrinking device space to accommodate the more complex RFFE. Historically, improvements in wireless system capacity have come from three main factors: network densification, more spectrum and greater radio link efficiency. Looking forward, these are still the controls we can use to continue to improve capacity. We can explain how these work using an extended form of Shannon's capacity equation:

$$R \leq C = \frac{m}{n} \cdot BW \cdot \log_2(1 + \frac{S}{N + I})$$

Equation 1.

The capacity (*C*) of a system is proportional to the channel bandwidth (*BW*) and the log of the Signal to Noise-and-Interference Ratio. *R* is the rate actually achieved. The parameter *m* is the number of independent spatial streams between the base station and the UE, and the parameter *n* is the number of users sharing the base station.[4]

In 3G and 4G radio systems, the channel BW is much greater than for 2G or older systems (up to 20MHz, compared to 200KHz or less), to increase capacity. But direct increasing of BW is no longer possible, as we will see below.

In the sections below, we summarize ways we can use these controls.

2.2.2. SPATIAL DENSIFICATION

By increasing the number of access points per area, the number of users per access point (*n*) decreases. This reduction improves the overall capacity of the wireless system.

More base stations per area also changes the requirements for the RFFE, reducing the transmit power required and reducing the needed sensitivity since the average distance from UE to BTS is reduced. We will see how that change affects the RFFE below.

2.2.3 MULTIPLE RX AND TX ANTENNAS

By adding multiple receive antennas on the user device, we can improve the system capacity in at least two ways.

First, using a second antenna for Receive (Rx) Diversity can (using one of several different techniques, like Selection Diversit, or Maximum Ratio Combining (MRC)) improve the Signal-to-Noise (SNR) ratio at the receiver and thus increase the maximum achievable data rate at the UE.

Second, implementing a multi-input multi-output system (MIMO, with multiple transmitter antennas at the base station) increases the number of independent signal paths (or spatial streams) from Access Point (AP) to UE, thus increasing parameter *m* and increasing capacity from AP to UE (downlink capacity).

---

[4] We will use the term "base station" generically to refer to eNodeB, base station, small cell, and Wifi access points.

Note that Rx Diversity requires 2 antennas, while MIMO needs 2, 4, 8 or even more antennas. Each additional antenna adds a potential additional spatial stream in the RFFE.

By adding multiple Transmit (Tx) antennas in the user device, we can improve the uplink system capacity in the same way as with multiple Rx antennas; that is, by improving SNR using diversity, or adding spatial streams using MIMO.

2.2.4 MULTIPLE SIMULTANEOUS FREQUENCIES

Earlier generation wireless systems use relatively narrow bandwidths (from 200 kHz) while 3G and 4G systems require wider bandwidths (up to 20 MHz). Despite being available in the standards, it has become very challenging to find more and more contiguous bandwidth available, especially in spectrum bands below 6 GHz. To increase bandwidth in order to increase capacity, newer systems have introduced the concept of multiple carriers (called Carrier Aggregation in LTE and Multi-Carrier in HSPA) where the user device receives or transmits on multiple frequencies at the same time. These multiple carriers can be in the same frequency band (intraband, either contiguous or non-contiguous) or in completely different frequency bands (inter-band). Multiple carriers can be added only for the downlink (Supplemental Downlink), or for both the uplink and downlink.

The multiple-carriers feature is especially attractive to wireless operators because of its ability to offer better utilization of discontinuous spectrum in different bands. The majority of wireless operators do not have access to a contiguous 100 MHz (or even a contiguous 20 MHz) of spectrum in all markets.

Note that the requirements for the components in an RFFE (for example, the PA) are affected by the bandwidth of each carrier or contiguous carrier-grouping, not by the number of carriers used in the whole RFFE. For example, a PA to carry a signal with a bandwidth of 20 MHz uses more current than one used for a bandwidth of 5 MHz (at the same average power). Multiple non-contiguous carriers instead affect the amount of hardware in the RFFE: more Power Amplifiers (PA), Low Noise Amplifiers (LNA), filters, and switches.

2.2.5. MORE SPECTRUM

As wireless systems have evolved and the number of users has increased, regulatory agencies have opened up new frequency bands to increase the total wireless capacity and meet user data demand. Plus, increasing data use per user means an increasing need for more bandwidth.

New licensed bands have opened up or are being opened[5]:

- At lower frequencies (700/800MHz and below) due to the switchover of analog broadcast television. However, this spectrum is not fully harmonized globally.
- At higher frequencies (2.3 GHz, 2.5 GHz and 3.5GHz)

Also, unlicensed bands (2.4GHz and 5GHz) have significant spectrum and are being used for mobile data use (offloading). While unlicensed, with no quality of service guaranties, they have the benefits of being harmonized, with significant contiguous bandwidth compared to some of the licensed bands.

---

[5] New regulatory approaches to allocate more bands, like Authorized Shared Assess (ASA)/ Licensed Shared Access (LSA), a concept to use under-utilized spectrum while increasing harmonization, has no direct impact on the RFFE other than adding a new band class.

Finally, discussions are ongoing about the use of very high frequency bands (e.g., millimeter-wave) for mobile broadband due to their large available bandwidths.

In order for an RFFE to use more bands it needs: more filters to cover the bands, more switches to direct signals among the filters, and PAs and LNAs to cover the new bands.

SPECTRUM FRAGMENTATION

More licensed spectrum is needed to meet mobile broadband demand, yet most large contiguous blocks of internationally harmonized spectrum have already been allocated and auctioned by many countries. Adding more spectrum has meant adding smaller non-contiguous pieces in a relatively ad-hoc manner with different regions having different blocks available, and different operators having discontinuous pieces of spectrum within a band. The worldwide spectrum for 3G and 4G is very fragmented.

Operators often own highly complex and fragmented spectrum assets obtained from multiple spectrum auctions. International roaming agreements demand support of additional frequency. The operators are faced with the problem of how to prioritize device support of frequency bands to enable seamless roaming over the whole operator's footprint. This usually leads to a complex market where multiple bands combinations are supported by different ranges of devices to keep RFFE complexity within manageable limits.

As we will see in Section 3, the methods of dealing with spectrum fragmentation include adding more hardware and more receive paths (Section 3.3) and using multiple simultaneous frequencies for both the receiver (Section 3.5) and transmitter (Section 3.4).

2.2.6. HIGHER-ORDER MODULATION

We can achieve higher capacity (in bits/second/Hz) for higher Signal to Noise Ratios by using higher order modulation. For example, instead of Quadrature Phase Shift Keying (QPSK), we can use 16 Quadrature Amplitude Modulation (QAM) or 64QAM. These higher order modulation schemes transmit more data in a given bandwidth (bringing R closer to C in Equation 1). However, they do not improve capacity without a cost, as we will see below. We see this increase in modulation order in both the Universal Mobile Telecommunication System (UMTS) and LTE systems.

2.2.7 DEVICE-TO-DEVICE COMMUNICATIONS & EMBMS

Device-to-Device Communication (D2D) allows devices near to each other to communicate without going through base stations. D2D can decrease latency and improve network capacity by offloading some data from the core network.

UE that support single-frequency TDD have a relatively easy means of supporting D2D; the UEs already have a shared transmit and receive frequency pair which is the same frequency. However, UEs that only support FDD become more complex—the transmit and receive frequencies need to be swapped in order to directly connect to another device.

For the RFFE, Enhanced Multimedia Broadcast Multicast Services (eMBMS) type services (point-to-multipoint) simply add a new receive band to the list of bands that a given UE may need to cover (if it is not already).

2.2.8. OTHER SIMULTANEOUS RADIOS

The multi-band and simultaneous operation with other non-cellular radio systems such as Bluetooth, WLAN, NFC, Frequency Modulation (FM), and/or Global Navigation Satellite System (GNSS) systems, are essential to current UEs. Those parallel and concurrent radio features add additional requirements on the RFFE design and optimization. The RF Front-End is required to operate over wide frequency ranges and, as more WWAN bands are added, they are often required to share resources (like antennas) for both WWAN and these other radios. All the UE RF components need to be ultra-linear, to prevent creating spurious energy in the receive bands of the other systems. These requirements will increase UE cost.

3. OPTIMIZATION OF THE RF FRONT-END

The proliferation of frequency bands, high-order MIMO and coexistence of multiple radios impose challenges to power and cost-efficient RF front-end design. The radio in a UE will be expected to support more than 40 WWAN bands in addition to WLAN, PAN, FM and GNSS. In this section, we will look at ways we can achieve these requirements and meet these challenges.

3.1 BRUTE FORCE VERSUS IDEAL

The RFFE in a UE can range from very simple (like those found in the original cellular phones) to the very complex (like in today's multi-band multi-mode cellular phones). As technology progresses, we seek to reduce (or hide) that complexity, moving towards the ultimately flexible Software Defined Radio. In this section, we briefly look at how this increase in complexity has occurred, and what the RFFE in an ideal radio would look like.

In Figure 2, we see the most basic configuration for a WAN radio, requiring only an antenna, duplexer, PA and an LNA. (A TDD radio needs a switch and filter instead of the duplexer.) An optional additional antenna, filter, and LNA allows Receive Diversity.

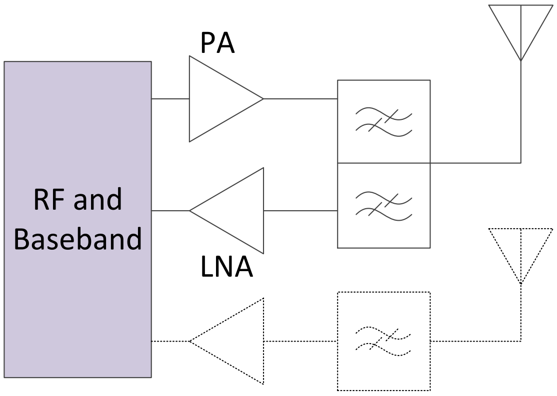

Figure 2. Basic WAN radio. Single band. Optionally, with diversity to increase coverage / capacity.

Later phones added coverage of a second band requiring more than double the RF hardware: an additional PA, LNA and duplexer, plus either another duplexer or dual-feed antennas (one feed for each band). Adding Receive Diversity means adding 2 more LNAs plus filters and either a diplexer, or another pair of dual-feed antennas.

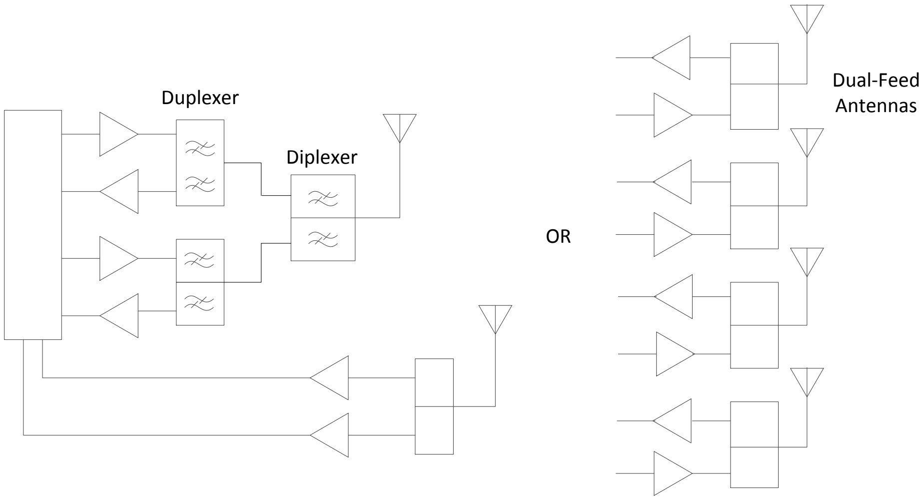

Figure 3. Second Generation Radio. Dual band. Optionally, with dual-feed antennas.

Adding more bands requires even more filters, LNA's, switches, and PAs for a single Multi-band Transceiver (MBTRX).

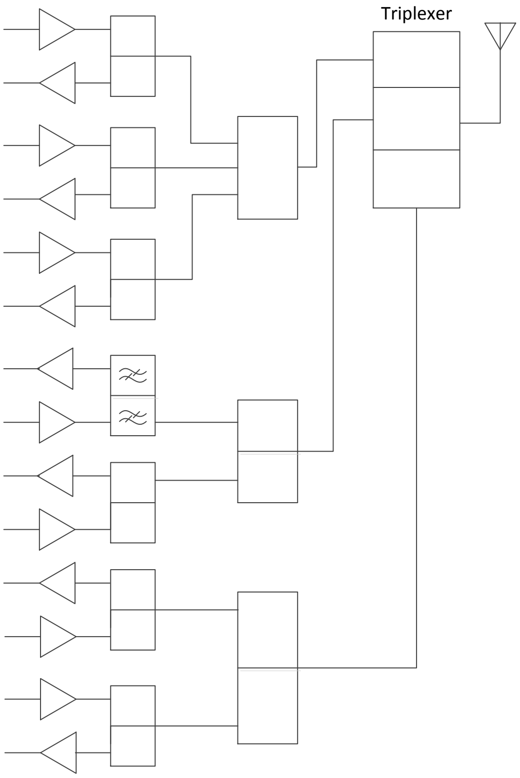

Figure 4. Multiband Radio (MBTRX). Antennas still potentially dual-feed

Adding MIMO to the UE means replicating all of the hardware in the RFFE for a MBTRX. Adding uplink or downlink CA means even more duplication of hardware.

UL-MIMO and Uplink Transmit Diversity require additional transmit paths (same frequency), while Multi-Carrier and Carrier Aggregation require additional transmit paths at different frequencies. Additional paths require additional components, such as filters, switches and power dividers, as well as longer routing. These additions result in a higher component count and higher loss in the RF Front End, which maps to higher cost and power consumption and lower receive sensitivity.

Clearly, the amount and complexity of the hardware is growing exponentially. This growth is unsustainable in the long run. Compare the above block diagrams with the “Ideal” 4x4 MIMO with Uplink/Downlink CA RFFE below.

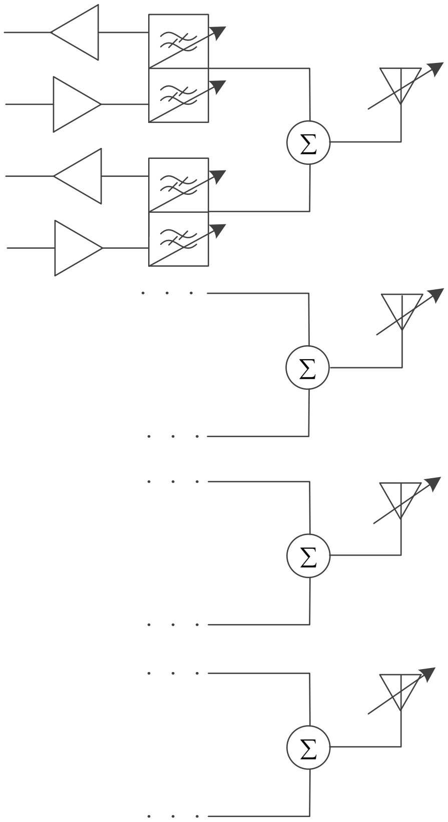

Figure 5. The Ideal RFFE with 4x4 MIMO and U/D CA. Four (or more) tunable TRx blocks with 2 (or more) simultaneous TRx sub-elements each for multiple simultaneous u/d links.

The Ideal RFFE would have at least 4 Transceiver (TRx) blocks capable of tuning across all the possible bands. Each TRX block would have at least 2 TRx sub-elements, allowing simultaneous transmission and reception at different channels (either in the same or different bands), on the same antenna. Each antenna would be tunable to adjust to different bands, including 2 simultaneously (to match the 2 TRs sub-elements), and to adjust for different external conditions (head/hand/object blocking).

Each component in the Ideal RFFE, even though it is tunable, still needs to meet the linearity and efficiency requirements of the individual components in a brute-force approach. But doing so in a tunable circuit is a significant challenge: the tunable elements themselves introduce nonlinearities or losses.

The overall goal of RFFE optimization is to move from the brute-force approach, towards the ideal approach while also adding to or increasing the functionality (more bands, more MIMO, more up and downlink carriers).

3.2 ANTENNAS AND ANTENNA TUNING

As more transceivers are added to the phone, there is an ever increasing need for more antennas:

- Receive / Transmit Diversity → 2x the number of Antennas
- Simultaneous Voice + Data (HSPA, LTE, …)[6], Dual-SIM → 2x or 3x
- More Bands → either more antennas (one for lowband, one for mid- and hi-band, ie. dual-feed), or broader bandwidth for each antenna
- MIMO → 2x, 4x, or even 8x

The screen size and battery size are also increasing, so the area allocated to RF components including antennas is shrinking. But the bandwidth (and frequency) of an antenna is proportional to its size, in particular, smaller antennas do not cover the lower bands well unless they are tuned to compensate for the decrease in size.

There are two methods of tuning antennas: Aperture Tuning, and Match Tuning. In Aperture Tuning, the antenna element itself is electrically adjusted.

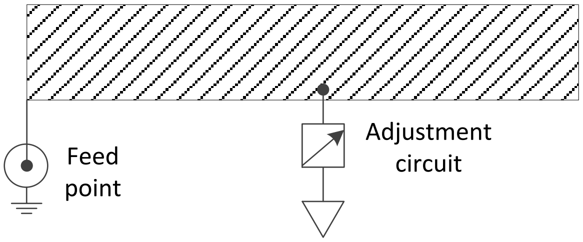

Figure 6. Aperture Tuning, where the behavior of the antenna itself is changed. The Feed Point is where the Antenna connects to the rest of the RFFE. The Adjustment Circuit can be of several different types, but will not be detailed here.

In Match Tuning, the RF match between the antenna and the rest of the RFFE is adjusted.

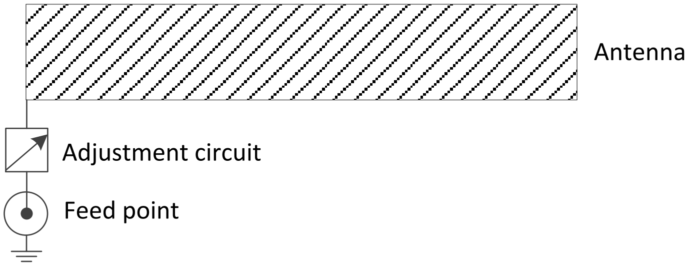

Figure 7. Match Tuning. The Adjustment Circuit changes the range of matched frequencies presented to the rest of the RFFE, but doesn't change the behavior of the Antenna itself.

---

[6] For example, SV-LTE prior to VoLTE

Each technique improves the match to the rest of the RFFE and therefore improves Transmit Efficiency.

Both Matching Tuning (MT) and Aperture Tuning (AT) techniques can be operated in 2 ways: either Open Loop, or Closed Loop. In Open Loop tuning, the Adjustment Circuit is varied based on which spectrum band(s) the UE is currently operating in. The adjustment rate is relatively slow.

In Closed Loop tuning, the Adjustment Circuit is varied based on both the band(s), and on real time measurements of the external environment. For example, if the UE is near the users head, the Antenna is detuned or mismatched. The UE can either directly measure the mismatch, or it can detect its position and change the Adjustment Circuit appropriately. Either way, the match to the rest of the RFFE is improved in a dynamic fashion.

Tuning of Antennas (and, in fact, any RF circuit) introduces issues of its own:

- More complexity. Adding active circuitry to each antenna (whether MT and AT) increases the complexity of the entire system and necessitates adding control signals back to the UE baseband.
- Distortion (due to non-linearity) of the Adjustment Circuit. The active circuitry in each antenna can cause distortions/harmonics to be created. Since the distortions are created in the antenna itself, additional filtering cannot remove them – the distortions must be of a small enough level to not degrade the performance of the receivers.

3.3 BAND SELECTION / DUPLEXING

Today's radio technology requires filtering in the RFFE for most modes of operation. Receive preselect filtering is needed to reduce the level of unwanted signals to the UE, and to allow optimization of the receiver's power consumption, size and cost. Transmit filtering is necessary to reduce unwanted emissions: a source of interference to other wireless services. UEs must support multiple radio services simultaneously (e.g., WWAN, WLAN, Bluetooth and GNSS), so the filters are important to mitigate self-interference as well as interference to and from nearby sources. Generally, each band and each path requires a separate set of filters.

The number of filters required by a modern handset is:

Total Number of filters = M x N

where;

M = Number of transmit & receive paths

N = Number of bands supported

Including more filters in a RFFE design add cost, area and complexity. They may cause more power consumption (due to increased RF losses on the Tx side), and poorer overall performance (due to RF losses on the Rx side) because of the need for RF switches/duplexer/diplexers between these filters.

3.3.1. MORE BANDS

The RFFE routes the signal from the PA or LNA to the antenna via the filters. The routing method depends on whether multiple bands need to be supported simultaneously. For non-simultaneous operation (i.e., TDD), a switch can be used. But, simultaneous operation (i.e., FDD) requires the bands to be duplexed (combined using bandpass filters) or diplexed (combined using highpass and lowpass filters).

An example of a RFFE supporting multiple bands is shown below. This example is for an RFFE capable of:

- RX Diversity
- 2 FDD Inter-Band Carrier Aggregation
- 3 FDD, Single Band or Intra-Band Carrier Aggregation
- 3 TDD, Single Band

Each path (RX or TX) is shown with separate amplifiers (LNA or PA). (In some cases amplifiers can be shared: this is not shown in the figure.)

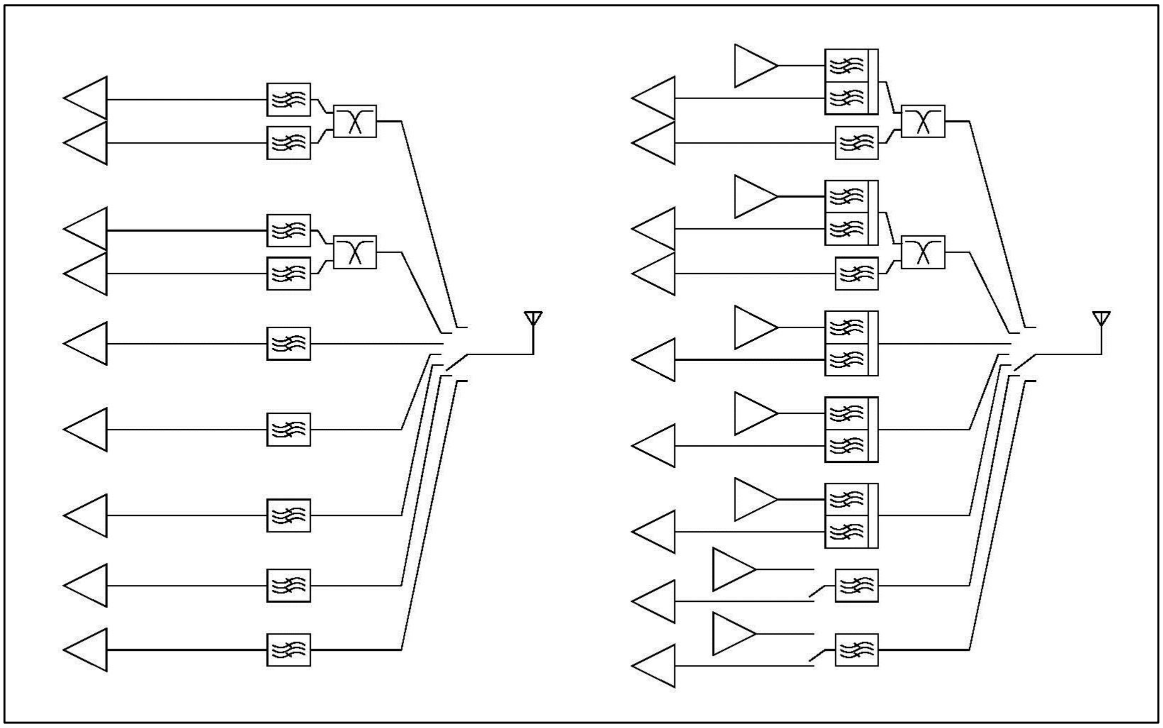

Figure 8. An example of an RFFE supporting multiple bands with Rx diversity, CA, and TDD.

As RFFE complexity increases, so does the loss from the additional switches, diplexers and routing. For example, an added loss of 3 dB, both degrades the receiver sensitivity by 3 dB, and forces the manufacturer to use a larger Power Amplifier. A larger PA results in higher power consumption (shorter talk/stand-by times), larger size (for thermal cooling) and higher cost—all negitive aspects directly observed by the consumer.

3.3.2. FEWER PRODUCT VARIATIONS

Because of the limitation on the number of bands that can be supported with today's RFFE technology (due to restrictions in board area and cost), the handset manufacturer is forced to produce and manage multiple product variations (i.e., SKUs). The desire to minimize the number of SKUs produced and managed results in not supporting under-utilized bands.[7]

3.3.3. RADIO CONFIGURABILITY

There are two possible approaches to addressing the challenge of having many bands while reducing product variations: either by switching between different filters, or by tuning. We discussed these two approaches in the "Brute Force versus Ideal" section above.

Currently, UE's address these challenges with advances in filter technology and packaging. Filters utilizing acoustic wave resonators (Surface Acoustic Wave (SAW), Bulk Acoustic Wave (BAW) and Film Bulk Acoustic Resonator (FBAR)) have made amazing progress. The filters have become smaller, cheaper and higher in performance. Although significant progress has been made and more is to come, a "Moore's Law" phenomenon cannot be counted on to address future handset needs. At some point, the RF losses in the RFFE limit the number of bands that can be supported. Packaging technology can help to address the multi-band challenge, by reducing size and (to a lesser extent) cost (driven by the cost of the individual devices), but ultimately RF loss will limit the number of bands supported.

A tunable filter, with the ability to cover the bands of interest, is an obvious solution. The past few years have shown many attempts at devices which could be made into such a filter. The technologies developed have focused on a tunable capacitor utilizing Microelectromechanical Systems (MEMS) and Ferro-Electric materials. Although significant progress has been made, the results have fallen short. It is not expected that these technologies have the inherent capabilities required (including Q and linearity) to produce tunable filters suitable for handsets. This area continues to be a priority for research; no key technology has been identified as promising yet. One reason is the performance/cost/size bar set by the acoustic filter industry is very high and the performance expectations have been incorporated into requirements written by industry and regulatory groups.

3.4 TRANSMITTER AND POWER AMPLIFIERS

The increasing number of bands and modes supported in a UE under similar space and cost constraints requires multi-band and multi-mode technologies, a highly integrated solution (for example, modules), and a low cost implementation (like Complementary Metal-Oxide Semiconductor (CMOS) PAs). The switching and multiplexing complexity of multi-band adds additional insertion loss after the power amplifier, while the higher Peak-to-Average (PAR) and high order of modulation (for 3G/4G and beyond) results in low PA efficiency (thereby decreasing talk-time). High efficiency PA Technologies like Envelop-Tracking (ET) and Digital Pre-Distortion (DPD) are starting to be attractive in mobile devices, while even more advanced PA technologies (like Doherty and Chireix – see Cripps, 1999) are being explored in both academic and industry research.

3.4.1. LOTS OF BANDS, LARGE FREQUENCY RANGE

Designing an RF Power Amplifier (PA) for mobile devices with good efficiency is a difficult challenge, but a critical one. The PA consumes more power than any other component in the entire RFFE, and directly affects device use and standby time. But the characteristics of the transistors in the PA change over frequency. Making an efficient power amplifier that works over a wider frequency range is an ever greater

---

[7] Under-utilization can be due to new bands, not fully deployed or populated and regional bands.

challenge. As more frequency bands are added at lower and at higher frequencies, the design of a single PA to cover every needed frequency becomes impossible. Typically, multiple PA's are added to a phone design to cover these higher and lower frequencies: for example, there might be one PA to cover 700-1000MHz (Low Band), another to cover 1700-2200MHz (Mid Band), and a third to cover 2300-2700MHz (Hi Band). Additional PAs would be needed for bands below 700MHz, at 3500MHz, or at 5GHz.

To mitigate the explosion of PAs required to cover all these bands, modules have been made to integrate multiple PAs into a single package. This integration helps reduce the total area and cost, and hide some of the complexity.

Carrier aggregation can actually leverage the band-splitting of the PAs by transmitting simultaneously in multiple bands.

MMMB (MULTI-MODE, MULTI-BAND) PA

A Power Amplifier Module (PAM) can include 2 to 4 discrete mode and band-specific PAs packaged together, while sharing the same control interface. A PAM can offer cost savings, boards area savings, simpler routing and reduced control complexity.

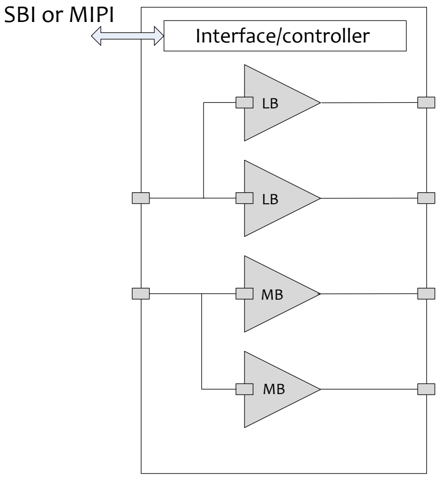

Figure 9. An example of a Power Amplifier Module.

Modern mobile devices often support 2G, 3G and 4G. As the number of bands and modes supported increases, converged multimode PA were developed. Converged multi-mode PA can support multi-standards (like UMTS, HSPA, CDMA, 1xEV DO and LTE FDD/TDD) in the same RF path. Some converged multi-mode PA designs even include Global System for Mobile Communications (GSM), General Packet Radio System (GPRS) and EDGE. However, due to differences in peak output power and modulation-type, these all-inclusive solutions have a trade-off in performance ("Jack of all-trades, master of none").

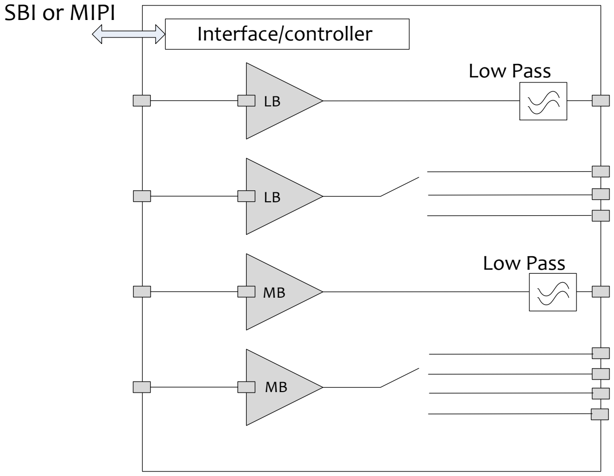

Figure 10. An example of a Hybrid Converged MMMB PA module.

A converged Multi-mode Multi-band (MMMB) PA Module reduces the total number of PA packages required, as well as associated passive components. It simplifies the board level routing with consolidated DC bias networks and control interfaces. Such MMMB PAs render comparable RF performance as discrete PAs in performance, while reducing part count and cost. Most MMMB PAs currently in the market are still a module-type solution with compound semiconductor Heterojunction Bipolar Transistor (HBT) PA and CMOS bias, interface and controller. The latest developments in CMOS PAs offer a unique opportunity for a fully integrated MMMB PA.

CMOS PA

Traditionally, compound semiconductor-based HBTs, have dominated the mobile power amplifier market. HBT-PAs provide high power efficiency and good linearity due to their intrinsic properties (they can handle higher voltages; have low on-resistance and a nearly insulating substrate). However, incorporating an HBT-PA into a MMMB PA module still requires CMOS circuitry for bias and mode control. Ideally, a MMMB PA module would include CMOS circuits only, allowing more integration and reduced size and cost.

A CMOS PA typically has limited high-power performance (due to lower drain voltage). In recent years, CMOS technology has made significant progress (in areas like high resistive substrate, silicon-on-insulator, through-substrate-via and low inductance flip-chip). These technology advances, combined with adaptive algorithms like Envelope Tracking (ET) and Digital Predistortion (DPD), have made CMOS power amplifiers more compelling.

CMOS provides a unique opportunity to integrate multiple PAs, RF switches, bias and mode control circuitry into single chip.

3.4.2. MORE TRANSMIT PATHS

Transmit Diversity (TXD) and Uplink MIMO (U-MIMO), as discussed above, are techniques to improve the Signal to Noise at the BTS or the number of simultaneous paths from the UE to the BTS. However, these techniques increase the number of PAs needed by 2x, 4x, or even more.

Little can be done to address this issue, other than putting more amplifiers in a single package (module), as discussed in the previous section. However, since these PAs will be used simultaneously, power dissipation becomes an issue, particularly since the module will have a lot of power in a relatively small area.

3.4.3 HIGHER CHANNEL BANDWIDTHS

The characteristics of the devices in the PA change over frequency. This change impacts not only the behavior from one frequency band to another, but also the behavior within a given RF channel when the channel bandwidth becomes relatively large. Designing a PA with good efficiency, while meeting distortion and emission limits, for a large channel bandwidth is a significant challenge. Note that unlicensed bands currently support wider channel bandwidths because they allow less stringent performance in other ways.

Research is currently underway in various venues to improve the performance of PAs while increasing their channel bandwidth and overall bandwidth. Presently, the poorer performance of the devices in these situations is included in the loss and power budgets for a given phone design.

3.4.4. HIGHER SIGNAL PEAK-TO-AVERAGE

A PA is designed to meet certain performance criteria like efficiency, linearity, and emissions. But the PA must be designed to meet these limits at the “peak” power. As the signal Peak to Average (P/A) is increased (to increase data rate), the average power that the PA can transmit necessarily decreases.

Thinking about this effect in the other way: if the signal P/A increases, the PA must be capable of handling even more power, even though the average power doesn’t change.

Two possible ways of addressing this efficiency loss include Digital Predistortion, and Envelope Tracking.

PA EFFICIENCY ENHANCEMENT: DIGITAL PREDISTORTION

The PA efficiency can be improved by driving the PA harder: forcing more output power even though the signal becomes more distorted and the SNR degrades. Digital pre-distortion (DPD) is a technique used to undo some of that distortion and improve the SNR. It has been used in base station design for many years. At the PA input the input RF signal is distorted in a way opposite to the way the PA itself distorts, resulting in low distortion output signal. The distortion parameters are decided based on PA characteristics and operating conditions. Sophisticated, low DC power consumption DSP techniques have improved, allowing adaptive digital pre-distortion for mobile device.

PA EFFICIENCY ENHANCEMENT: ENVELOPE TRACKING

Traditional PAs have fixed bias supply, while the instantaneous envelope of complex RF waveform varies with time. With a fixed bias voltage applied to a PA, PA efficiency is high when RF signal envelope is high and the PA is close to its saturation point. For signals less than the maximum, the PA efficiency is poor and a large amount of energy is lost in the PA itself. When the bias voltage can be lowered as the RF signal amplitude drops, less energy is used and average PA operating efficiency can be improved. There are several commonly used methods to adjusted PA bias voltage. Each requires different levels of sophistication in hardware and software.

Note, however, that the envelope tracker itself needs to be very power efficient, otherwise the overall power savings and increase in PA efficiency are lost.

There is ongoing research in the area of advanced PA's with higher efficiency.

3.4.5. FILTER / SWITCH LOSSES

The increase in bands means more filters and switches between the PA and the antenna. These filters and switches are losses, so the PA must transmit more power in order to get the same power to the antenna.

3.5 RECEIVERS / LNA'S

The enhancements to the wireless radio system impact the design of the Receiver in a UE, though much less so than the Transmitter. This section will address these design changes, and how their overall impact is mitigated.

3.5.1. LOTS OF BANDS, LARGE FREQUENCY RANGE

For the receiver, covering many bands over a large frequency range is less of an issue than for the power amplifier in the transmitter section. The increased power needed by wideband LNA designs is a small fraction of the total RFFE power.

3.5.2. MORE RX PATHS

The increasing number of LNAs needed for simultaneous operation (diversity / MIMO) does not increase the overall chip area by a great deal since the size of each LNA is relatively small. However, as the number of input pins/ports required increases, the LNA component size may need to increase (the component size may be limited by the number of pins, instead of by the active circuit / die area).

Complicated architectures, like Carrier Aggregation with both carriers in one band, mean even more receive paths that must be operated simultaneously.

3.5.3. HIGHER SIGNAL TO NOISE RATIO (S/N) FOR EACH CHAIN

More advanced radio systems (3G/4G) require a higher signal to noise ratio (SNR) to support higher data rates (see Equation 1). There may be an increase in power required for each LNA to allow the higher SNR. This power increase impacts the overall phone standby time rather than the talk-time. In talk-mode, the receiver power is a small fraction of the total phone power.

3.5.4. SIMULTANEOUS OPERATION

Operating several receivers and transmitters simultaneously in a phone (for example, with MIMO), can negatively impact one or more of the receivers in the phone. Careful filtering and circuit-board layout can help mitigate potential issues. But sometimes certain Transmitters or Receivers can directly impact a receiver by generating undesired signals in its band of operation. For example, the third harmonic of the LTE B17 Uplink falls within the B4 Downlink receive band.

3.5.5. MULTI RADIO TECHNOLOGIES IN EACH PHONE

For the receiver, the specific technology of the signal being received does not affect its design significantly. Rather, the issues are more related to band selection and duplexing, as discussed previously. Supporting 2G, 3G and 4G in a single receiver is not as difficult as doing the same for the transmitter. The only significant issue is the proliferation of LNAs/receivers for simultaneous operation.

3.5.6. DIGITAL / ANALOG RECEIVER ENHANCEMENTS

As discussed above, the most significant challenge for the receiver is the simultaneous operation of multiple transmitters and receivers. There are distortions generated that are very difficult to address in the usual ways (by increasing power, or adding more filtering). However, since the distortions are generated in a single device, knowledge of the sources of distortion can be used to mathematically (either using analog techniques or digital techniques) remove the distortion in the receiver which is being impacted.

The cost of such enhancements can be increased power consumption, increased complexity, or increased board/die area.

3.5.7. PAN / WLAN

Adding PAN and WLAN to a phone can negatively impact the design of the receiver. In the same way that simultaneous WAN transmitters and receivers can degrade the phone performance, so can simultaneous PAN / WLAN. Distortion signals can be generated which are difficult to remove even with filtering or increased power consumption. For example, if 2.4GHz WiFi signal couple into high band receivers, without adequate filtering the high signal level can overdrive and overwhelm the desired signal.

Care must be taken in chip and board design to minimize any impact. Additionally, much testing and debugging must be done to verify good performance for all the different RFFE components.

3.6 COMBINING FUNCTIONALITIES

It is a requirement for the RFFE in modern handsets to be small in size, low in cost and with minimal power consumption; thus being suitable for consumer electronic products. To achieve these goals, the industry is motivated to combine functionalities. The ultimate solution would be a set of components (PA, LNA, filter/duplexer, antenna) that could be configured for any band (e.g., 41 LTE bands defined by 3GPP) and the seven cellular modes (i.e., LTE-FDD, LTE-TDD, WCDMA, EV-DO, CDMA 1X, TD-SCDMA and GSM/EDGE). The industry has moved in this direction; examples are tunable antennas, multi-mode amplifiers (both PAs and LNAs). (See also, Section 3.4 on Multi-Mode Multi-Band PA modules.) The filters remain an enigma. No tunable filter technology is in production or under development that can match the performance, size and cost of acoustic filter technology (e.g. SAW, BAW, FBAR). Using filters that fall short of the established benchmarks (set by acoustic filters) result in larger size, higher cost and increase in power dissipation. Acoustic filters, switches and multi-mode/multi-band amplifiers provide the optimal solution today and the foreseeable future for the RFFE, as the industry continues to drive towards the ultimate solution.

To address the multi-band, multi-mode demands, while being constrained by the realities of available technology, manufacturers have developed advanced RF packaging technology for RFFE applications.

Advanced RF packaging includes innovations in packaging materials and techniques resulting in higher density of RF components, including 3D RF packaging which allows the stacking of devices.

An example of a basic RFFE module integrating a minimum of functionality is shown below. In this figure the module includes filters and duplexers supporting four FDD bands (including the RX Diversity path for each band) and the switches. The figure shows a set of filter external to the RFFE which could be used by the handset vendors to support bands with lesser economies of scale.

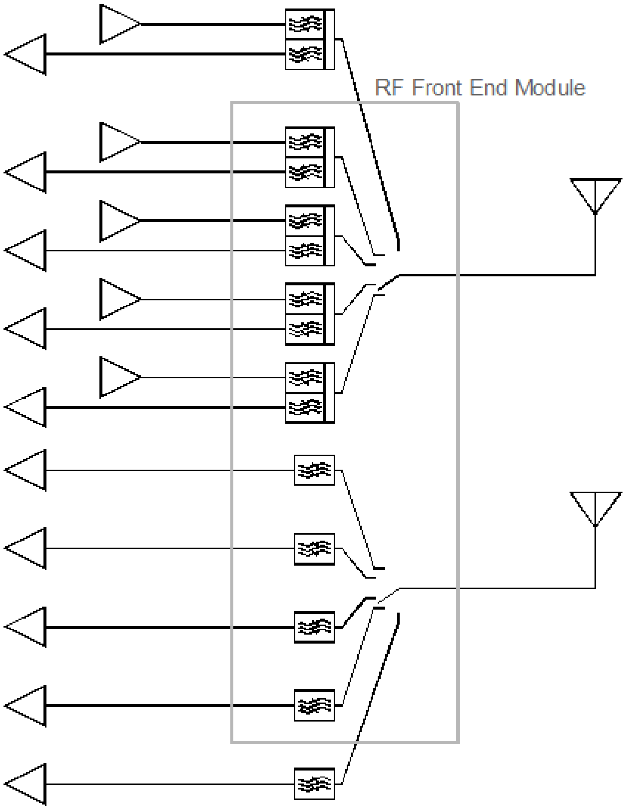

Figure 11. Example of a basic RFFE modue, including filters and diplexers supporting 4 FDD bands.

In contrast, an RF Front End Module which integrates all the RFFE functionalities is shown below. In this example the RFFE includes the following functionality:

- 5 FDD Bands, include RX Diversity
- 1 FDD Band with Intra-Band Carrier Aggregation, including RX Diversity
- 2 TDD Band, including RX Diversity
- Power Amplifiers
- Envelope Tracking Circuitry for PA
- Antenna Matching Circuitry
- LNAs
- Control Interface
- Example of amplifier sharing

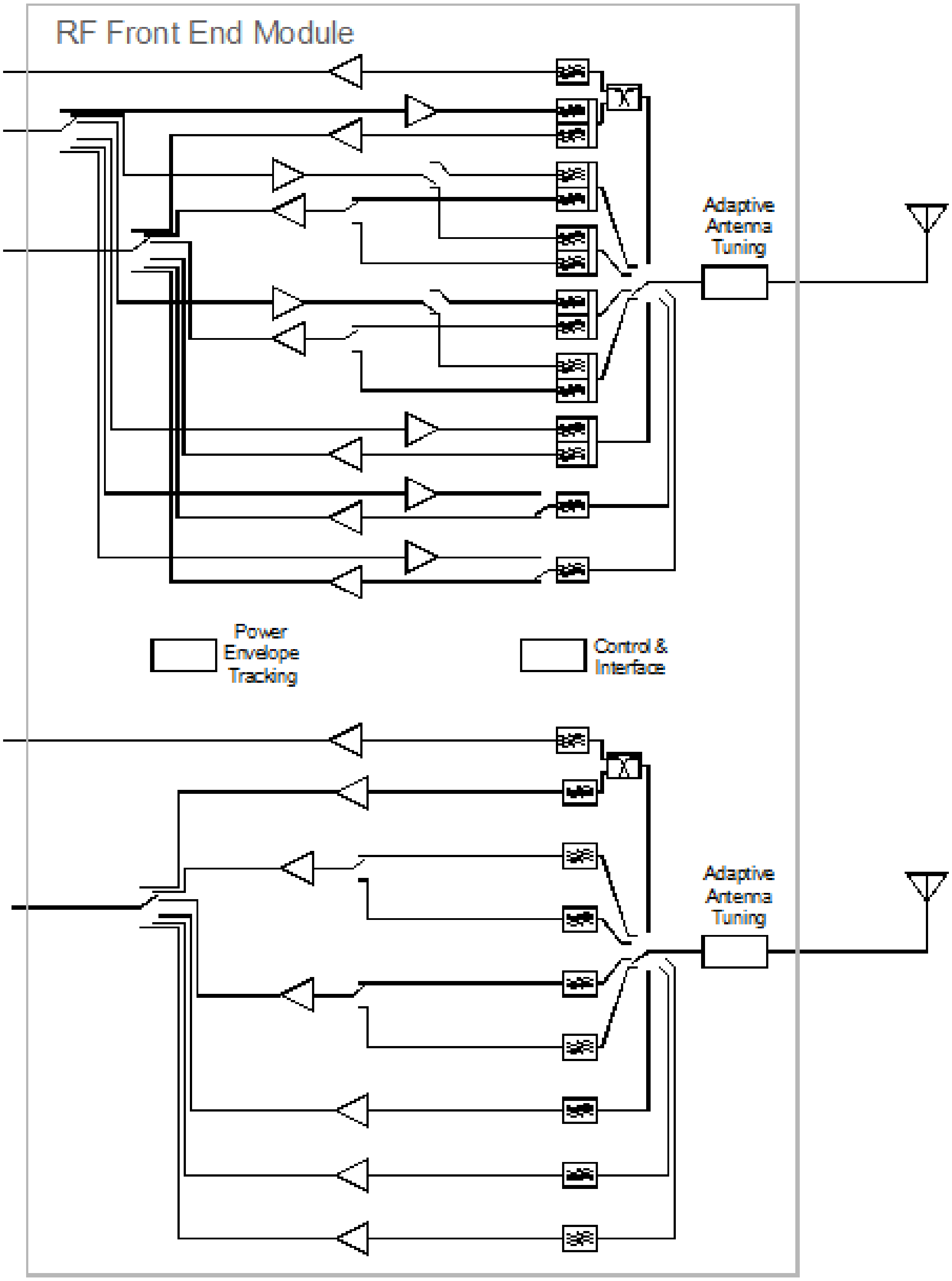

Figure 12. Example RFFE module with more functionality.

Although the RFFE for multi-band, multi-mode handsets are complex, RF Modules can hide this complexity and provide a simplified solution for the handset Original Equipment Manufacturer (OEM).

The RFFE footprint (i.e., area on the handsets printed circuit board) is reduced using an RFFE Module with advanced packaging techniques. Footprint reductions of up to 50 percent have been achieved, while maintain a low module height, which is critical for popular smartphones and tablets.

The RFFE Module and Radio chipset which have been pre-tested by the manufacturer provides a reduction in development time and cost allowing handset vendors to develop multi-mode/multi-band products faster and more efficiently.

RFFE Modules carry the burden of supporting bands which a customer may not use. For mainstream handset products in large scale markets the advantage is clear, however for markets with smaller economies of scale, the burden may become excessive. This burden is typically supporting bands not required by a customer or market (i.e. unused functionality). The optimal RFFE Module configuration(s) will ultimately be determined by the market.

3.7 OTHER ISSUES FOR THE RFFE

3.7.1. SELF-INTERFERENCE

As the number of radios in a UE increases, especially the number of simultaneously operating radios (like WWAN+WLAN, or Uplink CA, or uplink MIMO), the permutations and combinations of frequencies that can mix together (unintentionally) increases exponentially. Handset designers need to design for, and test, each of these possible sources of interference generated within a single UE. Thus, the design time and test time for new UE’s covering many bands has increased significantly, which raises the cost of design, development and test.

However, self-interference, as opposed to interference from nearby devices (co-existence), has one possible way of mitigating problems. The various radios can be coordinated or controlled at a higher layer to help prevent or at least mitigate the problems caused by self-interference. This mitigation, while helpful, is not without a cost: often throughput / maximum data rates suffer with higher-level coordination between radio systems: if one radio is off for a brief period, to prevent self-interference, the overall data rate must drop.

The ability to tune a radio can be seen as a way to reduce the size and complexity of multi-band radios. However, the active tunable elements themselves can generate harmonics and other distortion products that very negatively impact the self-interference problem. Non-tuned RFFE are passive and therefore not as likely to have these problems. Tunable elements in the Antenna are especially a problem since filtering to reject the harmonic / distortion is often not possible: the distortions fall within one of the receive bands.

3.7.2 HARDER CO-EXISTENCE

Co-existence refers to the situation where multiple UE-like devices are operated in close-proximity. In a similar way to the self-interference issue, the increase in receive / transmit bands in a UE makes the possibility of blocking and jamming from external signals more of an issue. Again, device manufacturers must design and test for each of the exponentially increasing number of possible combinations of external devices. For example, a user might have a tablet, phone, watch all operating close to a WLAN or small-cell BTS. Intermodulation / distortion generated in the RFFE can severely limit the performance of any or all of the devices, reducing their data rates or coverage range.

The limitations here are mitigated by increasing linearity in the RFFE (at the expense of area / power consumption). Unlike for self-interference, co-existence issues between devices are not readily fixable by coordination or another other higher-layer method.

3.7.3. LTE AND UMTS

From the RFFE perspective, only the ‘external’ characteristics of the signal matter. Examples for uplink are channel bandwidth, TX power range, PAR and emissions. Examples for the downlink are receive signal range and selectivity.

Most of the external characteristics of 3G and 4G signal align as a result of a conscious effort by the 3GPP organization. Although channel bandwidth differs, 3GPP standards define multi-carrier scenarios, thus 3G system can have aggregate signal bandwidths similar to 4G. Another difference is the PAR for the uplink signal, which impacts the handset’s PA. The uplink signal (for both 3G and 4G systems) is comprised of multiple individual channels, summed together to create an overall composite uplink waveform. Each individual’s channels modulation and power is independent. In reality, both 3G and 4G uplink waveforms have a range of PAR values, dependent on the active channels, modulation and power levels at any given time. Power Amplifier designers have been able to develop multi-mode PAs (i.e., both 3G & 4G), by adjusting the bias point (instantaneous or average depending of PA architecture) and impedance match.

Due to developments in PA technology, a common RFFE can be used by both 3G and 4G systems.

3.7.4. HIGH-ORDER MODULATION, DISTORTION AND LINEARITY

Moving to higher-order modulation (e.g., 64QAM) is the obvious response to the demand for higher data rates. Both 3G and 4G system utilize a range of modulation types dependent on the available RF channel, data rate required by the application and the robustness required by the individual channel. Both 3G and 4G system utilize high-order modulation. The RFFE’s role is to process the signal without introducing distortion sufficient to degrade the signal’s integrity. This drives the RFFE’s linearity requirement in both the RX and TX paths. Because both 3G and 4G systems utilize high-order modulation, the RFFE linearity requirements for 3G and 4G align.

3.7.5. TDD

TDD operation shares a single channel (on a time basis) for both uplink and downlink, where an FDD system has two separate channels separated in frequency, as shown in the figure below:

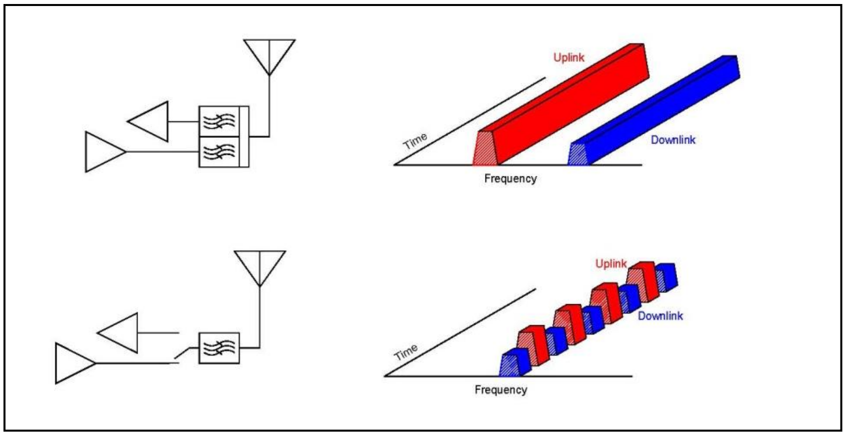

Figure 13. TDD versus FDD.

FDD systems require the uplink and downlink to be separated in frequency for isolation (i.e., no interference). TDD systems require the Uplink and Downlink to be separated in time for isolation (i.e., no interference). FDD & TDD operation is fundamentally incompatible within the same band, meaning they require separate bands when operating in the region.

From the RFFE perspective, TDD operation requires a single bandpass filter (shared by the TX and RX) and a switch, where the FDD operation requires a duplexer (two bandpass filters duplexer together).

The 3GPP organization identifies specific bands for FDD or TDD operation. The RFFE is configured to support the bands and modes (i.e., FDD or TDD) as required by the handset manufacturer.

4. CONCLUSION

The exponential grown of mobile data has created issues and opportunities in a key area of mobile devices, the RF Front-End. Consumers are using more mobile data in both the uplink and downlink, and demanding ubiquitous coverage on their devices, with bigger screens and batteries, leaving decreasing volume for the radio itself.

To meet these challenges, the wireless system and the RFFE in mobile devices must be optimized:

- There are more base stations in a given area (Spatial Densification)
- Each UE needs to have multiple antennas
- Each UE needs to operate simultaneously on multiple frequencies or channels, using more spectrum
- With higher order modulation, requiring a higher per-channel signal-to-noise ratio

Adding more bands and more RF features (MIMO, CA, Rx diversity, etc.), increases the complexity of the RFFE and requires optimization toward an ideal RFFE with more tunable elements, while meeting linearity and efficiency requirements. We attempt to suggest some ways to achieve these objectives either by switching between different filters or by tuning. Intensive research is ongoing in the areas of advanced filter technologies and packaging.

But the laws of physics and current technologies place limitations on what is achievable. Each receive or transmit chain needs at least some physical volume within the phone and uses some significant amount of power. There is an upper bound on the number of paths the RFFE can support. This upper bound is limited by size, cost and performance. This bound has already been reached by technologies available today and in the foreseeable future, causing UE manufacturers to produce multiple SKU's for a given device while keeping popular devices limited to large markets. Research is underway to push these boundaries further, but the design and optimization of the RFFE is still a significant challenge.

5. REFERENCES / RELATED WORKS

1. Network Densification: the dominant theme for wireless evolution of 5G. http://ieeexplore.ieee.org/xpl/articleDetails.jsp?reload=true&arnumber=6736747

2. 4G Mobile Broadband Evolution: Release 10, Release 11 and Beyond, 4G Americas, February 2014

3. Meeting the 1000x Challenge: The Need for Spectrum, Technology and Policy, 4G Americas, October 2013

4. Christopher Taylor, “Mostly commonly used linear bands in top-selling smartphones”, RF&Wireless Components Service, March 2014

5. Nick Cheng; J.P. Young, “Challenges and requirements of multimode multiband power amplifiers for mobile applications”, IEEE CSICS, 2011

6. Peter Asbeck; Lawrence Larson; Don Kimball; J. Buckwalter, “CMOS handset power amplifiers: directions for the future”, IEEE CICC, 2012

7. Ray Pengelly; Zhancang Wang; Damon Holmes; Mustafa Acar; Robin Wesson; Mark van Heijden, “Modern high efficiency amplifier design: envelope tracking, doherty and outphasing techniques”, Microwave Journal, April 14, 2014

8. Lawrence Larson; Donald Kimball; Peter Asbeck; Paul Draxler; Junxiong Deng; Ming li, “Digital predistortion techniques for linearized power amplifiers”, Proceedings of Asia-Paccific Microwave Conference, 2006

9. Cripps, 1999. RF Power Amplifiers for Wireless Communications. Artech House.

10. Mobile Experts, 2012. Handset RF Front Ends.

6. ACKNOWLEDGEMENTS

The mission of 4G Americas is to advocate for and foster the advancement and full capabilities of the 3GPP family of mobile broadband technologies, including LTE-Advanced, throughout the ecosystem's networks, services, applications and wirelessly connected devices in the Americas. 4G Americas' Board of Governors members include Alcatel-Lucent, América Móvil, AT&T, Cable & Wireless, Cisco, CommScope, Entel, Ericsson, HP, Intel, Mavenir Systems, Nokia, Openwave Mobility, Qualcomm, Rogers, Sprint, T-Mobile USA and Telefónica.

4G Americas would like to recognize the significant project leadership and important contributions of Anne-Lise Thieblemont of Qualcomm, as well as representatives from the other member companies on 4G Americas' Board of Governors who participated in the development of this white paper.