Thin-film, bulk-wave acoustic resonators (FBARs) work along the same principle as SAWs -- electrical energy is converted to mechanical energy -- but unlike SAW devices, the energy is directed into the bulk. The primary mode of sound energy is longitudinal, whereas in a SAW the sound energy is either a Rayleigh wave or a surface-skimming wave.

On a superficial level it would appear that FBAR technology must be inherently superior to SAW technology. Bulk devices have better power handling abilities, and the bulk device is fairly insensitive to surface contamination and surface adsorbates. Furthermore, having the electrical fields contained between the two electrodes guarantees minimum coupling of electrical fields with outside metal surfaces and capacitance being determined by the spacing, area and dielectric constant of the piezoelectric.

That said, the relative ease of manufacturing of SAWs, using one or two photoresist and metal deposition steps, made them the technology of choice.

The FBAR device suffers from the fact that a large acoustic impedance mismatch must exist on both sides of the piezoelectric. This requires either the inclusion of a Bragg reflector on the bottom side of the piezoelectric or an air/crystal interface under the resonator. The Bragg reflector must be made up of at least 3 to 4 pairs of acoustically-dissimilar materials [1].

If one chooses to make an air/crystal interface on both sides of the piezoelectric some non-standard IC processing technique is needed. One obvious path is to etch silicon from the back side. Either a thin p++ silicon layer or a Si3N4 thin film material can be used as an etch stop. Fig. 1 shows an early device (circa 1993) using a low-stress silicon nitride layer as an etch stop [2].

![Fig. 1: Early (1993) FBAR Using SiN Layer As Etch Stop](image)

One obvious drawback to this technique is manufacturability. Creating etch holes through the back of the supporting structure (silicon) weakens it -- making it especially susceptible to breakage. Another
drawback when using wet etches for silicon is that the walls are set at a 54.7° angle, thus the actual area used for each resonator and or device becomes large, reducing the number of die per wafer.

Surface-machining techniques would eliminate the need for etching from the backside altogether. One approach using a tapered sacrificial layer has been described [3]. A second approach [4] is to deposit a sacrificial glass layer (which can subsequently be removed in HF), open up holes for structural supports and pads and deposit CVD tungsten into the openings. The deposited tungsten is then polished down to the sacrificial layer leaving behind tungsten plugs of various cross sectional areas.

Fig. 2 shows an FBAR resonator using this technology. In both Figs. 1 & 2 electrical heaters were placed around each of the resonators for temperature compensation techniques.

Process Design

The current Agilent process uses free-standing membranes that are anchored at the edges to the silicon. Rather than etch from the backside, shallow pits are formed on the front side of the silicon and back-filled with a low temperature oxide. The wafer is then polished back to the silicon surface, leaving a single smooth coplanar surface alternating between oxide and silicon. Wafers processed to this point are referred to as “starting material.” At this point all further processing is done at a relatively low temperature [5].

The deposition and patterning of the electrodes and AlN piezoelectric use standard IC processing tools and techniques. After the electrodes, piezoelectric, mass loading (to move the center frequency of certain resonators a few percent lower relative to others), and pads, the low temperature oxide is etched in a dilute HF solution (DHF). “Weep holes” are used at the corners of the resonator to allow
the DHF access to the sacrificial oxide. This step is referred to as the “release” step. Fig. 3 is an SEM profile of a finished FBAR cleaved near one edge of the “swimming pool.”

The electrode material used for Agilent FBARs is molybdenum [4]. The requirements and restrictions posed by the process greatly limited the choice and among some of the boundary conditions are low acoustic attenuation, high electrical conductivity, and process compatibility with AlN. This latter requires that there be etch selectivity between electrode material and AlN. Finally, the metal must be robust to the final “release” step.

One of the biggest problems with free-standing resonators is parasitic lateral modes. These modes give rise to “suck out” in and around the passband response of the filter. Lateral parasitic modes set up resonances defined by any two parallel edges of the resonator. By “apodizing” in the resonator layout [6] we can eliminate most of these modes. First, no two edges are allowed to be parallel. Second, where possible, pentagons are used to minimize any reflections in the x-y plane.

Fig. 4 shows the overlay of two filters (laid out side by side), with and without apodization.

![Fig. 4: Micrograph Superimposing Two 5-GHz Filters With And Without Apodization (Vert. Axes 5-dB Per Division, Horiz. Axes 50-MHz Per Division; Span: 5.0 To 5.5 GHz)](image)

**Reliability**

Reliability requirements differ significantly from application to application. For wireless applications the critical issues involve mechanical stress (vibration, mechanical shock, and dropping), environmental factors (temperature + humidity, ESD, and temperature cycling), and operating life (high, low and wet high-temperature operating life). Extensive reliability studies were performed on FBARs to demonstrate that it is an intrinsically robust and reliable device. Early results showed that our FBAR filters can withstand up to 4 W (+36 dBm) input power [7]. One filter design demonstrated the ability to handle 10 W of input power with no damage.
High temperature operating life tests at 1 W input power and 100°C showed no failures after 1500 hours. Stress conditions were largely derived from the JEDEC standards and listed below in Table 1. All parts were tested before and at intervals during the stress tests with no failures observed in any test.

<table>
<thead>
<tr>
<th>STRESS</th>
<th>CONDITION</th>
<th>SAMPLE SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet High Temperature Storage</td>
<td>85°C/85% R.H.</td>
<td>50</td>
</tr>
<tr>
<td>Temperature Cycle</td>
<td>JESD22-A104-B Condition G</td>
<td>50</td>
</tr>
<tr>
<td>Thermal Shock</td>
<td>JESD22-A106-A Condition C</td>
<td>50</td>
</tr>
<tr>
<td>High Temperature Storage Life</td>
<td>JESD22-A103-A</td>
<td>50</td>
</tr>
<tr>
<td>RF High Temperature Operation Life</td>
<td>$T_x P_{in} = 30$ dBm, 100°C</td>
<td>16</td>
</tr>
<tr>
<td>RF Low Temperature Operation Life</td>
<td>$T_x P_{in} = 30$ dBm, -40°C</td>
<td>16</td>
</tr>
<tr>
<td>RF Wet High Temperature Operating Life</td>
<td>$T_x P_{in} = 30$ dBm, 50°C/95% R.H.</td>
<td>16</td>
</tr>
<tr>
<td>Electrostatic Discharge</td>
<td>HBM: JESD22-A114-B</td>
<td>20 per model</td>
</tr>
<tr>
<td></td>
<td>MM: JESD22-A115-A</td>
<td></td>
</tr>
<tr>
<td>Mechanical Shock</td>
<td>JESD22-B104-A</td>
<td>10</td>
</tr>
<tr>
<td>Vibration</td>
<td>JESD22-B103-A</td>
<td>10</td>
</tr>
<tr>
<td>Drop Test</td>
<td>Height: 152 cm</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 1: Summary Of Reliability Tests

Mean-time-to-failure (MTTF) lifetimes under CDMA-use conditions (this implies variable power levels) were determined to be 35 years with 90% confidence and 90% reliability. For a constant input power of +27 dBm at 25°C, MTTF is predicted to be 13 years with 90% confidence and 90% reliability. Lifetimes were determined using input power on the transmit unit as the acceleration factor. Input power was applied at the frequency where the power absorption in the resonator is at a maximum (this is typically around the –4dB point on the high frequency side of the passband for the transmit unit. This frequency is also outside the passband for PCS). While the power was being absorbed, the negative coefficient of frequency caused the filter response to shift down in frequency. A computer-controlled feedback loop was therefore used to automatically keep the input power frequency at the –4 dB point at all times during the test guaranteeing that the power is being absorbed.
by the filter rather than being reflected or passing through the load. Real-time computer monitoring of the filter response was used to determine when a failure occurred. Actual filter use would not normally be at these frequencies. Hence the MTTF calculation are a worse case lower limit. Actual MTTFs should be considerably higher.

A total of 88 units were tested at various input powers from 1 to 4 W with 83 of 88 units being stressed to failure. Failures generally occurred as a result of electrical overstress at the leading input edge of the series or shunt resonator. The failure distribution was modeled using the Weibull probability density function (PDF) and the Inverse Power Law accelerated life model (Eq. 1). The power coefficient (n) was calculated to be 54 and the shape parameter (β) was determined to be 0.6. A shape parameter of less than 1 implies a declining failure rate with failures occurring due to infant mortality.

\[
MTTF = \frac{1}{KV^n} \Gamma \left( \frac{1}{\beta} + 1 \right)
\]

\[
PDF = \beta KV^n (KV^n t)^{\beta-1} \frac{1}{e^{(KV^n t)^{\beta}}} 
\]

Eq. 1

where: K, n = constants
V = stress level (power)
β = shape parameter
Γ = gamma function

Model Results:

\[\beta = 0.6\]

\[K = 1.12e-84\]

\[n = 54\]

Lifetimes were calculated using 2 acceleration factors: a heavy phone user factor (2000 min/month) and a suburban CDMA cumulative damage factor. The cumulative damage model accounts for the integrated damage from the various powers of a typical suburban CDMA use model shown in Fig. 5. The suburban CDMA use model is based on the phone being used at various powers over time. This model more accurately reflects the conditions that the user will see rather than assuming a single

![Fig. 6: Progression Of HBM ESD Improvement From Dec. 2000 To March 2001](image-url)
constant power level. For example, the phone will be on approximately 1% of the time at 1 W, 2% of the time at 800 mW, and 1% of the time at 630 mW in this use model. The suburban use model requires higher power usage than an urban use model due to fewer, farther apart bay stations, thus providing a more stringent test condition.

ESD measurements were mainly done using the Human Body Model (HBM) and the Machine Models (MM). HBM testing involves charging the device under test with a high voltage pulse across a capacitor (100 pF) and through a 1500 Ω resistor which simulates the resistance of the human body. Fig. 6 shows the improvement in average HBM filter tolerance as a function of time. This was realized through a number of design and process improvements. For comparison we measured a commercial 1.9-GHz SAW filter, which had an HBM ESD sensitivity of 200 V.

Modeling

There are two types of models used to describe FBAR resonators and filters. Both types of modeling are used extensively for device design, optimization, layout and sensitivity analysis.

The first is based on the Mason model [8,9]. This is a “physical” model and uses: dielectric constant, mass densities, stiffness coefficients, coefficients from the piezoelectric stress tensor and thicknesses of the physical layers. This model calculates the fundamental frequency of the resonator as well as the effective $k_t^2$ of the device, extracting “$k_t^2$ effective” from measurements taken from the Q-circle.

The intrinsic $k_t^2$ (for the longitudinal mode) is defined by the ratio of the square of the piezoelectric stress tensor, $e_{33}$, divided by the stiffness coefficient $c_{33}$ (in the presence of an electric field) and the dielectric constant $\varepsilon$. However, in a filter, it is the difference in frequencies of the parallel and series resonance (fp,fs) of the Q circle.

In the limit of infinitesimally thin electrodes, $k_t^2$ is equal to $\pi/2 * (fs/fp) / \tan[(fs/fp)]$ [10]. For small values of $k_t^2$, one can replace this equation with “$k_t^2$ effective” = $4.8*(fp-fs)/(fp+fs)$ [2]. Note, this approximation will give “$k_t^2$ effective” greater than the theoretical limit for the intrinsic $k_t^2$ (6.5% [11]) for certain range of electrode thicknesses. This is both predicted by the Mason model and borne out in experiment.
The second model [12] is referred to as the modified Butterworth-Van Dyke (MBVD) model and uses a simple electrical model, Fig. 7, to predict the resonator response. The MBVD model is especially useful for modeling filter responses built from modeled resonators. The MBVD model fits well with current commercial circuit simulators[13] where optimizing code can be run to achieve the desired filter response and sensitivity analysis can be done to find where the yield “knobs” are.

Fig. 8 shows a measurement of a resonator and a “best-squares” fit to that resonator. The fitted values for each element is; \( C_p = 1.8 \, \text{pF} \), \( G_s = 1.18 \, \text{Mhos} \), \( R_{\text{series}} = 1.02 \, \Omega \), \( L_m = 79.4 \, \text{nH} \), \( C_m = 80 \, \text{fF} \), and \( R_m = 0.65 \, \Omega \). Far from resonance, the resonator looks like a plate capacitor. At the frequency \( f = \frac{1}{2 \pi \sqrt{L_m C_m}} \), there is a series resonance \((f = f_s)\) with a real impedance around 1 to 2 \( \Omega \). At \( f = \frac{1}{2 \pi \sqrt{L_m C_m C_p/(C_m+C_p)}} \), there is a parallel resonance \((f = f_p)\) with a real impedance around 1500 \( \Omega \) to 4000 \( \Omega \). In a filter using a half-ladder topology, the shunt resonators are mass loaded such that \( f_s \) and \( f_p \) are 1% to 3% lower in frequency than the non-mass loaded resonators used for the series elements.

Typically, the \( k^2 Q_p \) product -- in manufacturing -- ranges from 30 to 40, where the effective \( k^2 \) is taken from \( f_s \) and \( f_p \) (the two frequency points where the Q-circle, Fig. 9, crosses the real axes). \( Q_p \) is extracted from \( R_p \), the maximum resistance as measured off of the Q circle at \( f_p \). Fig. 9 shows a Q-circle of a more recent R&D design where an \( R_p \) of 4200 \( \Omega \) was measured. The \( k^2 Q_p \) for this resonator is \(~95\). Typically, \( Q_s \), the series resonant Q, is very high 1400 to 2000. \( Q_p \) is typically 50% to 80% of \( Q_s \). New design changes have “leveled the playing field”. Now, \( Q_s \) and \( Q_p \) are the same and both are very high. Some improvement in lateral mode suppression is needed prior to production release. Fig. 10 shows the raw filter performance (S11 and S12) using these resonators.
Duplexer performance strongly depends on the Q and kt2 of the raw filters. However, more is needed for successful productization. We use external inductors [14] to distribute attenuation-poles across the passband of the other filter. This ensures the rejection and isolation meet the stringent specifications required of commercial CDMA duplexers. Monte Carlo simulations are used for sensitivity analysis highlighting the effects of variations in $kt^2$, Q, impedance, external inductors etc. [15].
**Performance In The Market**

There are two major schemes implementing wireless phones emerging in the market place. TDMA is a half-duplex operation that depends on a switch to separate transmission and reception. This “walkie-talkie” type operation is used in GSM. The other scheme, CDMA, uses full-duplex operation in the phone. This requires the receive chain to “listen” to incoming signals as low as –106dBm input power, while the Tx chain is “talking” at powers up to 30 dBm. This simultaneous action is done in two frequency bands separated by 20 MHz in the 1920 MHz PCS bands (~1%).

Until recently, ceramic duplexeres were the only viable technology that could meet the stringent requirements for front end filtering. SAWs do not have sufficient Figure of Merit (kt²*Qp product) to

![Graph showing performance](image-url)

meet the electrical requirements for a commercial duplexer.
The physical size of the ceramic duplexer poses two problems for the handset manufacturer. The first is the actual area taken up by the duplexer. There is a continuing quest for more functionality in the phone (e.g., GPS for E-911), but the PCB is limited in size. The second, and perhaps just important as the area, is the height. Current duplexers are 4 to 5 mm high. The FBAR duplexer is under 2 mm. The handset manufacturer would like to have all of the components at the same height. “Manhatten” profiles are undesirable as it means either greatly increasing the thickness of the phone (creating “dead” space inside the phone, or a need for specially shaped plastics for either the case or the battery to handle the necessary cut out for the duplexer. The Sony C-Z100SPR wireless phone is an example of where the plastics for the battery were specially designed to fit around a ceramic duplexer.

In contrast, Sanyo has introduced the new Sanyo SCP-6000, currently the thinnest, lightest CDMA phone on the market. This phone uses the Agilent FBAR duplexer.

Much work has been done since the first publication on FBAR duplexers in 1999 [16]. At that time, the roll-off of the Rx was inadequate to meet current specifications for a wireless handset. Fig. 11 shows the Tx and Rx response. Furthermore, this duplexer used die on board assembly for proof of concept and the part was not hermetic. The next two years was spent developing a hermetic package and improving specifications to be usable in commercial products.
Figs. 12 and 13 show a more recent duplexer (assembled in Malaysia) filter response and the isolation between the Rx port and the Tx port. Note that the isolation is 54 dB in the Tx band and 44 dB in the Rx band.

Future Work

Agilent is currently working on ways to take advantage of the fact that FBAR resonators are manufactured on silicon. One extension is to leverage micromachining techniques to make wafer scale chip level package. Fig. 14 shows a recent working Tx full-band interstage filter in a “microcap” package. The top wafer (or microcap wafer) is bonded to the FBAR wafer and the individual packaged filters are singulated using conventional sawing techniques [17]. The packaging is done in the same IC fab that the FBAR wafer is made [18]. Currently, wirebonding is done through the microcap lid. But, future plans will incorporate the vias into the lid.

References