In most electronic materials the change in dielectric constant with applied electric field is an effect too small to be useful or even easily measurable. But for a special class of high-permittivity materials the effect can be quite pronounced. This chapter explores some challenges and opportunities for exploiting tunable dielectrics to make reconfigurable, adaptive, frequency-agile RF devices. We will focus on thin-film barium strontium titanate (BST) materials, discuss design and modeling of BST varactors, and survey some circuit implementations and heterogeneous integration efforts.

4.1 Introduction

BST Ceramics

The field-dependent permittivity (i.e. the “tunability”) of high-permittivity dielectrics has been known for quite some time, and its potential for use in RF circuits was recognized immediately [1]-[7]. Barium-titanate, BaTiO$_3$ (abbreviated as BTO) and related compounds are now considered the prototypical “high-κ” or high dielectric constant materials for this purpose. In bulk ceramic form, barium-titanate is a ferroelectric at room temperature, with a ferroelectric-paraelectric transition at $T_c \sim 116^\circ$C, the so-called Curie temperature. Figure 4-1 illustrates the behavior of a ferroelectric material around the Curie temperature. In the ferroelectric phase below $T_c$ the material exhibits memory effects or hysteresis in the polarization-field response that can
be exploited for non-volatile embedded memories [8]. Above $T_c$, the material is *paraelectric* with a very large dielectric constant and field-dependent nonlinearity. In both phases the electrical properties are strongly temperature-dependent. Although the paraelectric phase is of most interest in this work, we still colloquially refer to the material as a ferroelectric.

The Curie temperature in bulk BTO can be easily manipulated by mixing with other materials or compounds to allow for room-temperature operation. Figure 4-2a shows the variation in Curie point with various additives [9]. Strontium titanate, SrTiO$_3$ (STO) is an interesting and useful choice because it also has a high permittivity, so the dielectric constant and nonlinearity (tunability) remain high as the Ba/Sr ratio is changed, but the Curie temperature decreases almost linearly with the amount of strontium as shown. The solid solution of BTO and STO is barium-strontium titanate, Ba$_x$Sr$_{1-x}$TiO$_3$ (BST), with $x$ specifying the mole fraction of barium. Three representative examples are marked in Figure 4-2a. The Curie temperature falls below room temperature when $x < 0.7$.

![Figure 4-2](image.png)

**Figure 4-2** – (a) Variation in Curie temperature of bulk BTO with various additives (after [9]). (b) Structure of BST showing oxygen octahedra surrounding the titanium ion (after [10]). (c) Field dependence near the Curie temperature for a high-barium BST ceramic (after [2]).

BST has a cubic perovskite structure, with the large barium and strontium ions occupying the corners of the unit cell as shown in Figure 4-2b [10]. The oxygen ions form an octahedral “cage” surrounding the small central titanium atom. In the paraelectric phase, the high-permittivity of the material derives from the fact that the titanium ion can be easily displaced by an applied field, yielding a large induced dipole moment or polarization. This effect is further enhanced by long-range ordering effects at low temperatures. The displacement or polarization (indicated by arrows in the figure, assuming a vertical field) begins to saturate at high fields, leading to a reduction in the small-signal effective dielectric constant. The measured dielectric response for a representative bulk ceramic Ba$_{0.73}$Sr$_{0.27}$TiO$_3$ mixture is shown in Figure 4-2c [2].

Bob York
Significant changes in the dielectric constant with applied field are apparent in Figure 4-2c, especially near the Curie temperature, but the concurrent strong variation with temperature in bulk BST material is a source of concern for practical applications. Another limitation associated with tunability in bulk ceramics is the rather high voltage required to achieve the field strength needed for significant tunability (typically 10-20 kV/cm or higher), and associated breakdown considerations. For these and other technological reasons the field-dependence of tunable dielectrics was not widely exploited, and the effect remained a laboratory curiosity for decades. However, it is important to note that BTO and derivatives have been extensively used in high capacitance-density passive components, enjoying huge commercial success in both the ceramic capacitor and thermistor industries. Thus considerable resources have been and continue to be employed towards understanding and improving these materials for commercial use.

**Thin-Film vs. Thick Film**

In the early 1990’s, two separate technological developments – the discovery of high-temperature superconductors, and evolutionary scaling issues in Silicon Dynamic Random Access Memories (DRAMs) – revived interest in BST and STO for RF applications. In the case of high-temperature superconductors, low-loss tunable dielectrics were sought to integrate with the superconducting circuits to create high-performance filters [11]. For the CMOS and DRAM industry, high-κ alternatives to SiO₂ were sought to maintain the steady rate of improvements predicted by Moore’s law [12-13]. In DRAMs a high dielectric constant is desired to reduce the size of the storage cell capacitor; the capacitor dielectric is deposited in very thin films (<100nm) and must be process-compatible with the remaining CMOS transistor circuitry. Significant resources were dedicated to developing candidate dielectrics such as BST for this demanding application.

For these new applications, methods were developed and/or refined for depositing thin-film BST by industry-standard and high-throughput techniques such as RF sputtering and Metal-organic Chemical Vapor deposition (MOCVD), and the electrical properties of the thin-films proved to be significantly different than bulk ceramics of similar composition. As
shown in Figure 4-3, not only are the dielectric constants much lower, but more importantly
the temperature dependence is drastically different, with no obvious ferroelectric-paraelectric
transition and much smaller temperature coefficients. In addition the tunability (change in
dielectric constant with voltage) is quite large, aided in part by the higher fields that can be
sustained in thin films in comparison to bulk materials. Naturally the voltages required to
achieve a given field are also much lower than for bulk materials.

These remarkable differences between thin-film and bulk BST are generally favorable as
far as potential RF applications are concerned, resolving many of the limitations that
discouraged early application efforts. As these properties became more widely appreciated,
several high-frequency devices and circuits using thin-film STO and BST were reported [15]-
[22] (these are just a few representative references on thin-film devices only, not thick-film or
bulk ceramics). Many of these early reports also used high-temperature superconductors,
reflecting the evolutionary origins of the work.

In the last decade, steady progress has continued towards an understanding of BST thin-
films. It is now believed, for example, that the large differences observed in thin-film BST in
comparison to bulk materials are probably due to a combination of factors including so-called
“size-effects” [23] and residual strain in the materials arising from high-temperature growth
on thermally mismatched substrates [24]-[26]. The relevant material deposition technology
has also evolved continuously, and some progress has been made in understanding the failure
mechanisms and long-term reliability of the materials (more on that later).

On the other hand, it is also now widely appreciated that many factors can influence the
behavior of thin-film BST, including the method of deposition, growth temperature,
composition, substrate, film thickness, processing conditions, and contact metallurgy. This
chapter will focus, therefore, on behavioral or phenomenological modeling approaches that
capture the general features of tunable dielectric devices, independent of the particulars of the
technological implementation.

**Potential for RF applications**

Thin-film BST has already been commercially exploited for high-capacitance density
decoupling capacitors [27]-[29], but the focus of attention here is on voltage tunable
integrated capacitors (varactors). Two questions arise: where can BST varactors be useful,
and will the devices offer some compelling advantage(s) over alternative technologies?

![Figure 4-4 – Comparison of diode and dielectric varactors characteristics. BST varactors have
no forward conduction region, and hence can sustain large RF voltage swings, especially near
zero volts (exaggerated here for illustrative purposes).](image-url)
Semicontductor diode varactors have been exploited for decades [30], and a large body of work exists on the design of varactors and varactors circuits. Tunable filters, phase-shifter, linearization networks, and voltage-controlled oscillators are several important examples. Despite the different physical mechanisms responsible for the tuning, diode varactors and dielectric varactors have similar tuning ratios and quality-factors. Dielectric varactors do have one key advantage relative to diodes: there is no forward conduction region. As shown in Figure 4-4, this allows for improved power handling and simpler biasing compared with diodes. Dielectric varactors also require less sophisticated processing for a given Q-factor and operating frequency, and can be easily integrated with other high-Q passive components, factors that tend to lower overall implementation cost. On the other hand, diode varactors are an established and mature technology, and at present it is unclear whether the linearity and cost advantages of dielectric varactors will be sufficient to displace diode varactors.

Another emerging technology for wireless applications is RF Micro-Electro-Mechanical systems (MEMS). MEMS devices seem best suited to a switching function, although analog varactors have been realized using MEMS techniques. In comparison to MEMS, dielectric varactors have the following advantages: they are physically smaller; require lower control voltages; require less processing and less complicated packaging (lower cost); have a fast intrinsic response time; and most importantly have excellent RF power-handling characteristics, allowing for hot-switching without degradation. The only real negative in relation to MEMs is the nonlinearity of the device, but this can be managed with techniques discussed later.

So relative to diode varactors and MEMS, there is at least a potential opportunity for dielectric varactors in applications requiring large RF voltage swings and low cost. The analog front-end of mobile wireless communication devices is one such application, and might potentially benefit from frequency agile filters, matching networks, tunable antennas, and phase-shifters that can be realized with tunable dielectrics.

Doping, Composites, and Other Tunable Dielectric Materials

Most work in tunable dielectrics for RF applications has centered on BST and STO because of the unique combination of high tunability and good RF loss characteristics, but there are other candidate materials [31]. Many are structurally similar perovskite structures of the form ABO₃, where the B-site cation is typically small in comparison to the A-site cation and often one of the following three: 1) titanium, e.g. lead titanate, PbTiO₃ (PTO) and lead zirconium titanate, Pb(ZrₓTi₁₋ₓ)O₃ (PZT); 2) niobium, e.g. lithium niobate, LiNbO₃, and potassium niobate, KNbO₃ (KNO); or 3) tantalum, e.g. potassium tantalate, KTaO₃ (KTO). These materials are either ferroelectric (like BTO) or incipient ferroelectrics (like STO), displaying high permittivity and tunability and similar temperature dependencies. PTO and PZT are technologically important materials for piezoelectric and embedded memory applications, but the RF losses are typically high in comparison to BST [32,33], limiting their usefulness in varactor applications at high frequencies. As with BTO, the addition of STO to lead titanate results in a solid solution (Pb,Sr)TiO₃ (PST) with a Curie temperature that depends on the Pb/Sr ratio [34]. It has recently been shown that strontium-enriched PZT, e.g. Pb₁₋ₓSrₓZr₀.₅₂Ti₀.₄₈O₃ (PSZT) has improved characteristics as compared with PST [35]. Similarly K(Ta,Nb)O₃, a solid-solution of KTO and KNO, has been shown to have potentially useful properties for RF applications [31]. Clearly there are some common threads of scientific thought at work in the development of these materials.

Another emerging material worthy of special mention is bismuth zinc niobate [36], Bi₁.₅Zn₁.₀Nb₁.₅O₇ (BZN). In contrast to BST and many of the other materials mentioned
earlier, BZN is a non-ferroelectric material with a cubic pyrochlore structure, exhibiting a relatively large dielectric constant (150-200) and a low loss tangent (< $10^{-4}$ at 1 MHz). BZN thin films show a significant field-dependence of the permittivity, with more than 2:1 change in the dielectric constant at field strengths of ~2.4 MV/cm [36,38]. Although BZN bulk ceramics exhibit a dielectric relaxation and high losses at microwave frequencies [37], we have recently shown that dielectric losses of thin-film BZN capacitors remain comparatively low at least up to 20 GHz [38-39]. Hence BZN thin films appear to be attractive for microwave tunable applications [40]-[42].

In most cases of practical interest the thin-film materials will be polycrystalline, with defect and grain-boundary structures that depend on growth temperature, choice of substrate and electrodes, and impurity concentrations. It is well known from the ceramic capacitor industry that oxygen vacancies are a problem in these materials, often linked with leakage, reduced lifetime, and bias-induced performance degradation. A time-honored solution to these problems is the addition of small amounts of compensational “dopants” [43]; this has been shown to also work well with thin-film STO [44] and BST [45]-[48]. Since the precise amount of compensation doping is difficult to ascertain a priori, amphoteric dopants such as yttrium [43,49,50] and erbium [51] appear to be a promising solution.

From a fundamental standpoint high losses are an unavoidable companion of high permittivity and tunability, and hence there is always an inherent tradeoff between loss and tunability. Another approach to manipulating the properties of tunable dielectrics adds rather large amounts of non-tunable compounds, usually a low-loss/low-permittivity linear dielectric (e.g. MgO). The resulting “composite” material has a reduced dielectric loss tangent, but of course this comes at the expense of reduced tunability, which as we will see is fundamentally linked to the zero-field permittivity. It can be shown by relatively simple electrostatic considerations [52] that one can never increase the figure-of-merit (often defined as a product of tunability and Q-factor) of a tunable dielectric using such composite mixtures. Composites are also more prone to structural defects and contamination, making them undesirable for a robust and reproducible high-volume thin-film manufacturing process. Fortunately there are more effective techniques to address dielectric losses and linearity concerns, discussed later.

**Scope of this work**

No materials have yet demonstrated characteristics superior to BST when all of the critical variables are considered. However, as noted earlier, BST is itself a complex material with electrical properties that are dependent on many physical factors including Ba/Sr ratio, grain size, dopants, film thickness, temperature, frequency, and field strength. BST films are sensitive to the methods used in material preparation and device processing. Electrical properties of BST devices are also strongly dependent on the substrates and contact metallurgy. For these reasons and in combination with its technological importance, BST is sometimes referred to as the “BeaST” of electroceramics.

Our goal here is to develop a simple but general framework for modeling tunable dielectric devices and circuits with respect to important independent variables (frequency, applied electric field, film thickness, and electrode size) in a way that is transparent to the technological details. The approach is largely phenomenological in order to obtain a simple closed-form modeling of these electrical properties. We will focus exclusively on thin-film, parallel-plate capacitors, since these are the most important structures from a practical standpoint. No attempt will be made to understand or model the observed properties of the materials or devices from a detailed consideration of the underlying physics. The reader is
referred to [31] and [53] and references therein for a more in-depth overview of materials, deposition, and processing technology associated with tunable dielectrics.

4.2 Low-Frequency Measurements and Modeling

Device characterization at RF and microwave frequencies (>100MHz) is complicated by several factors that stem from the high-permittivity of the films, as well as difficulties inherent in making accurate high-frequency measurements on high-Q reactive components. Low-frequency (<100 MHz) measurements can therefore be especially useful for basic material characterization of tunable dielectrics, and much of what has been learned about BST devices is based on low-frequency characterization techniques.

Test Structure Design and Impedance Analyzer Measurements

At low-frequencies, relatively simple large-area test structures can be used and fabricated with a minimal amount of processing, increasing the turnaround time between material growth and characterization, an important consideration for circuit development. With carefully chosen device sizes and frequencies, certain parasitics (such as electrode resistance and inductance) can be safely neglected to simplifying the data analysis. Most importantly, highly accurate measurements can also be made over several orders of magnitude in frequency using balanced-bridge I-V methods [54], and such broadband dielectric spectroscopy is especially helpful for developing accurate circuit models [55]. The Agilent 4294A impedance analyzer is one example of an instrument that uses this technique, can be easily configured for on-wafer probe measurements.

A useful test structure for thin-film BST characterization is shown in Figure 4-5. After depositing the BST film on a metalized substrate, the BST film is subsequently patterned to form a “mesa”, and a second metal contact layer is deposited on the film and bottom electrode simultaneously. Since the BST films are usually quite thin (100-500nm), the variation in height between the top and side contacts is negligible, thus the structure can be
directly probed by GSG (ground-signal-ground) coplanar probes as shown, provided the top contact area is larger than the probe tips. In practice, devices as small as 20 $\mu m \times 20 \mu m$ have been easily characterized using this method. Although simple two-pronged (GS) probes could also be used, a GSG test structure has the advantage that the series resistance arising from the bottom electrode is reduced by a factor of two (more on that later). Furthermore, GSG probes have been shown to yield more accurate calibration and measurements.

A variety of substrates, contact metals, and process conditions can be used and these will affect the results in various ways. In much of our work the BST films for low-frequency characterization were deposited by RF magnetron sputtering onto platinum-coated (100-200nm) c-plane sapphire substrates. The BST films were then wet-etched using a buffered HF solution, and Pt(100nm)-Au(200nm) top contacts were then deposited by a lift-off procedure. More details on the processing can be found in [26] and [56].

Figure 4-6 – (a) Typical capacitance-frequency and (b) Q-frequency data from 0-10V in 1V increments. Data taken on 160nm Ba$_{0.5}$Sr$_{0.5}$TiO$_3$ films using an Agilent 4294A Impedance analyzer with a 200mV test signal amplitude).
Representative data on a 160nm thick, nominally stoichiometric 50/50 BST film (Ba$_{0.5}$Sr$_{0.5}$TiO$_3$) is shown in Figure 4-6. This data was recorded using an Agilent 4294A analyzer and RF probe station; the capacitance data assumes a parallel $G-C$ ( admittance) model for the device, and the $Q$-factor (inverse loss tangent) is computed as $\text{Im}\{Y\}/\text{Re}\{Y\}$. The frequency was swept from 40Hz to 110MHz (the limits of the instrument) at different DC bias voltages from 0-10V in 1V increments. The instantaneous catastrophic breakdown voltage for this particular film was >22V, corresponding to a field of >1.3 MV/cm.

Clearly there is a complicated dependence on frequency and bias for both the capacitance and the $Q$-factor (loss). The following sections will develop a fairly complete phenomenological model for these dependencies. It is important to note that the particular device measured in Figure 4-6 was a large-area device, nominally $2.5 \times 10^4 \mu$m$^2$, such that peripheral effects (discussed later) are expected to be negligible. It is also important to mention that the impedance analyzer calculates device admittance according to $Y = \Delta I / \Delta V$, where $\Delta I$ is the complex AC current (in-phase and quadrature component) that flows upon the application of an AC voltage $\Delta V$ [54]. If the AC voltage is small enough, the data approximates the desired small-signal admittance $Y = \partial I / \partial V$. The choice of oscillation amplitude is thus dependent on how strongly the material properties change with voltage, which in turn depends on material composition, film thickness, etc. If the oscillation amplitude is too large, the calculated capacitance will differ from the small-signal value, and the loss will appear to increase as a result of nonlinear frequency conversion (harmonic generation) and self-heating effects. In the data set of Figure 4-6, a 200mV oscillation amplitude was used.

**Capacitance-Voltage Relationship**

In Figure 4-6a we can see that at any one frequency the capacitance decreases monotonically as the bias field is increased. For a symmetrical device (identical top and bottom contacts) the result is independent of the polarity of the bias field. At zero field the capacitance starts at some maximum value $C_{\text{max}}$ (dependent on electrode area, film thickness, frequency, temperature, etc.). As the applied DC field increases the small-signal capacitance $dQ/dV$ decreases monotonically. At some voltage $V_r$ the capacitance is reduced to $C_{\text{min}}$; we define the tunability, $\tau$, as the ratio of maximum-to-minimum capacitance at this voltage

$$\tau \equiv \frac{C_{\text{max}}}{C_{\text{min}}}$$  \hspace{1cm} (4.1)

The tunability thus defined is dependent on the choice of $V_r$. In this analysis, $V_r$ can be chosen arbitrarily, but it is shown later that the “2:1” voltage $V_2$ is an obvious choice.

Recently we have derived a simple closed-form expression for the $C(V)$ relation [57], beginning with a power-series expansion for the field-polarization relation of the form [1,14,24]

$$E = \alpha_1(T)D + \alpha_2D^2 + \ldots$$  \hspace{1cm} (4.2)
where $\alpha_i(T)$ is the inverse of the zero-bias permittivity (temperature-dependent) and $\alpha_3$ describes the nonlinearity of the material. In the context of ferroelectric films (4.2) is called the Landau-Devonshire-Ginzburg (LDG) model. Irrespective of any physical justification for this model, there is a simple phenomenological basis: the even symmetry of the $C(V)$ relationship for dielectric varactors means that there can only be odd terms in a power-series expansion for $E(D)$, and higher order terms can be neglected because dielectric breakdown is usually encountered before these terms become significant.

For an ideal capacitor (no interfacial layers or space-charge) we can assume that the $E$ -field and flux density $D$ are uniform throughout the film, and relate to the external applied voltage and charge through

$$E = V / d \quad D = Q / A$$

(4.3)

where $d$ is the capacitor thickness and $A$ is the area. This transforms (4.2) into

$$V = \frac{\alpha_i d}{A} Q + \alpha_3 \frac{d}{A} Q^3$$

(4.4)

The small-signal capacitance is defined by

$$C(V) = \frac{dQ}{dV}$$

(4.5)

Using (4.4)-(4.5) and the definition of tunability in (4.1), we showed [57] that the capacitance nonlinearity can be modeled as

$$C(V) = \frac{C_{\text{max}}}{2 \cosh \left[ \frac{2}{3} \sinh^{-1} \left( \frac{2V}{V_2} \right) \right] - 1}$$

(4.6)

where

$$V_2 = \frac{4V_1}{(\tau + 2)\sqrt{\tau - 1}}$$

(4.7)

is the “2:1” voltage at which $C(V_2) = C_{\text{max}} / 2$, an easily measured quantity. Experimentally there are only two parameters that define the ideal $C(V)$ curve: $C_{\text{max}}$ and $V_2$. Once we know $V_2$ for a given device, (4.6) can also be used to determine the voltage required to achieve a desired tunability.

Although the explicit $C(V)$ relation in (4.6) has only recently been derived, the formula for the 2:1 voltage (4.7) has been known for some time [1]; in fact this relationship defines an implicit $C-V$ curve, since using (4.1) we can show that

$$V = \frac{1}{4} V_2 \left( \frac{C_{\text{max}}}{C(V)} + 2 \right) \sqrt{C_{\text{max}} - 1}$$

(4.8)

This is quite useful for predicting the voltage at which a certain capacitance value is reached, an issue relevant to the control circuit design in frequency-agile networks. For completeness we also note that the original LDG expansion (4.2) has now been successfully inverted in a simple closed form

$$D(E) = \frac{3}{2} \varepsilon_0 E_2 \sinh \left[ \frac{1}{3} \sinh^{-1} \left( \frac{2E}{E_2} \right) \right]$$

(4.9)
where \( E_z = \frac{V_z}{d} \) is the field at which the permittivity changes is reduced by a factor of 2, and \( \varepsilon_0 = \frac{1}{\alpha} \) is the zero-bias permittivity. This is a potentially useful result for EM simulators.

The model presented here compares quite favorably with data on large area devices (large area-to-periphery ratios) where edge effects can be neglected. Figure 4-8 shows a representative example, with excellent agreement out to voltages in excess of \( 8V_z \). This is somewhat remarkable given that there are only two parameters in the model.

**Effect of Interfaces and Breakdown on Tunability**

The field-dependent tunability is ultimately linked to the inherent nonlinearity of the material (the third-order term in (4.2)) and not surprisingly the tunability is dependent on material composition and deposition conditions. However, the tunability is also affected by the electrode-BST interface, and also depends indirectly on the breakdown field, both of which lead to a thickness-dependent tunability.

![Figure 4-9](image1.png)  
Figure 4-9 – (a) Thickness dependence of the effective dielectric constant calculated from zero-bias capacitance data at 1 MHz for various film compositions (sapphire substrates and Pt electrodes). (b) Corresponding plot of inverse capacitance density [26].

Figure 4-9 shows the measured dielectric constants (calculated from the zero-bias capacitance data \( C_{max} \) at 1 MHz) as a function of film thickness for three different material compositions on sapphire substrates. For very thick films the apparent dielectric constant saturates; this can be explained by the presence of a non-tunable interfacial capacitance...
[24,58] associated with the electrode-BST interfaces. These are sometimes called “dead” layers since they do not respond to the field in the same manner as the remainder of the film. There is no general agreement yet on the exact origin of this capacitance, but possible candidates are: 1) stiffening of the material’s optical phonon mode near the interface [59]; 2) an effective interfacial capacitance due to field penetration into the electrodes [60]; or 3) an interfacial capacitance associated with near-surface charge traps [61]. In any case, the interfaces are modeled as suggested in Figure 4-10 by a fixed capacitance $C_i$ in series with the remaining bulk film capacitance $C_b$ such that

\[
\frac{1}{C(V)} = \frac{1}{C_i} + \frac{1}{C_b(V)}
\]  

(4.10)

Using the parallel-plate capacitor formula we can eliminate the dependence on electrode area and write

\[
\frac{1}{\varepsilon(V)} = \frac{t}{\varepsilon_i} + \frac{d-t}{\varepsilon_i(V)}
\]  

(4.11)

indicating that a plot of the inverse dielectric constant versus film thickness should approach a straight line for $d \gg t$, and the resulting y-intercept gives the interfacial capacitance density (the slope yields the bulk permittivity $\varepsilon_b$). From Figure 4-9b we see that interfacial capacitances are typically on the order of 40-60 F/μm². Note that this effect may be present in any thin-film capacitor, but for low-permittivity materials there is a negligible impact because the interfacial capacitance is so large in comparison to the bulk capacitance density. On the contrary, for high-permittivity materials the bulk and interfacial capacitance densities can be comparable in value, particularly for very thin materials that are desired for low control voltages, and thus the effect of the dead-layers can be significant.

Using (4.10) we have shown [57] that the composite $C(V)$ relationship has the same functional form derived earlier, but with thickness-dependent parameters $C_{\text{max}}(d)$ and $V_2(d)$. For device optimization it is helpful to make the thickness dependence more explicit, especially with regards to the tradeoffs between control voltage and power handling or linearity. Once the interfacial capacitance density is determined, we need only measure the maximum capacitance and 2:1 voltage at some nominal material thickness $d_0$, then the general thickness-dependent tuning parameters become

\[
C_{\text{max}}(d) = \frac{1}{\frac{1}{C_i} + \frac{d}{d_0} \left[ \frac{1}{C_{\text{max}}(d_0)} - \frac{1}{C_i} \right]}
\]

\[
\frac{V_2(d)}{V_2(d_0)} = \sqrt{\frac{d_0}{d}} \left( \frac{C_{\text{max}}(d_0)}{C_{\text{max}}(d)} \right)
\]

(4.12)

The $C(V,d)$ functional is then uniquely determined for any film thickness by the specification of $C_i$, $C_{\text{max}}(d_0)$, and $V_2(d_0)$.

Bob York
Figure 4-11 shows the data for three devices of identical electrode area, processed from three different thicknesses of low-barium BST with Pt electrodes. From a thickness series an interfacial capacitance density of ~32 fF/μm² and a bulk permittivity of ε_b = 420 were computed. The data was extracted from broadband RF data taken on a network analyzer as described later. Using these parameters, and using the 575nm material as the reference, the theoretical curves for each device were generated from (4.6) and (4.12). Excellent agreement is observed using the dead-layer model for the thickness dependence.

Figure 4-12 is an example of the dependence of tunability on deposition conditions, in this case the Argon/Oxygen gas mixture during sputter deposition (see [62] for details). It has been shown [63-65] that varying the gas mixture and background pressure influences the stoichiometry of the film, and this in turn has a strong influence on the zero-bias permittivity ε_b. A key observation is that regardless of the initial dielectric constant, all curves tend to approach a similar asymptotic permittivity at high fields. Thus the tunability is not only determined by the zero-field permittivity, it is also dependent on the maximum field that can be sustained by the material.

The breakdown field E_{br} governs the maximum voltage V_{br} ≈ E_{br}d that can be applied to the material. In simple experiments, breakdown is usually taken to mean an instantaneous, catastrophic device failure; thus published C-V curves like those in Figure 4-12 are usually terminated at a field or voltage just below breakdown. Using this definition, the dielectric strength of thin-film BST is typically 1.5-2 MV/cm. Although this number can vary somewhat with thickness, deposition and processing conditions, it is generally consistent with the general downward trend in dielectric strength observed in many materials as the dielectric constant increases (Figure 4-13).
The next logical question that arises is what maximum field can be safely applied to the materials for long periods of time. It is well known that the lifetime of thin-film capacitors is degraded by prolonged exposure to high-fields [67]; in high-\( \kappa \) ceramics this is due to many effects such as accumulated charge injection, migration of charged defects (such as oxygen vacancies), self-heating, etc. [23,68]. The safe operating fields for tunable RF devices can therefore only be determined by careful, long-term reliability studies. A small but growing body of literature exists on this subject for BST (e.g. [69]-[74]), and although there is no clear consensus, it is believed that bias fields should be kept below 500-600 kV/cm for long-term operation at temperatures up to 85°C.

Thus if we constrain the field strength to some maximum value based on reliability considerations (typically well below the catastrophic failure field), then the maximum voltage and hence tunability becomes thickness dependent. This is illustrated in Figure 4-14 for some representative field strengths, using the data from Figure 4-11 and the model developed in (4.12). For example, an application requiring 2:1 tunability at a maximum field of 500kV/cm would require a 250nm film using this particular material. Since the control voltage also depends on thickness, this example highlights a tradeoff between the control voltage and tunability as a result of both breakdown and interfacial capacitance considerations. This is important in some applications where low control voltages are desirable, such as battery-operated devices.

It is interesting to note that the dielectric strength of thin-film BST is quite a bit larger than what is often quoted for bulk ceramics of similar composition. A simple empirical model for thickness-dependent breakdown in bulk materials is given in [75] as

\[
E_{\text{br}} \approx 0.988(d[\mu \text{m}])^{-0.39} \text{[MV/cm]} \tag{4.13}
\]

where the fit was determined from samples on the order of a few millimeters thickness. It is almost absurd to expect this formula to hold for thin-film materials that are 4 orders of magnitude thinner, but in fact the predictions are in reasonable agreement with measurements on high-quality samples such as those shown in Figure 4-12. Assuming a similar relationship applies to the maximum fields determined from long-term reliability studies, then we can conclude that thin-film materials will always have higher tunabilities than bulk.
materials of similar composition, because they can sustain much larger fields. This observation is consistent with most of the results reported in the literature.

**Low-frequency Loss and Dispersion**

Turning attention back to the measured Q-factor in Figure 4-6b, we can identify some important trends with frequency and voltage: At the lower-frequencies the Q-factor decreases with applied field, approaching a linear frequency dependence at moderate to large bias fields; In the middle of the frequency range the Q-factor is initially frequency independent at zero bias, and then increases slightly with bias field; At the upper end of the frequency range the Q-factor begins to roll off but also increases with bias field.

![Parallel conductance versus frequency at different bias voltages](image)

Figure 4-15 – Parallel conductance versus frequency at different bias voltages, computed from (4.14) and the data in Figure 4-6.

Below 10MHz the behavior can be understood using a parallel conductance model, where the conductance $G$ can be related to the measured $Q$ in Figure 4-6b as

$$G = \frac{\omega C}{Q} = \omega C \tan \delta$$

(4.14)

This is shown in Figure 4-15. At very low frequencies the conductance is frequency-independent (constant) under bias, increasing significantly with voltage; this is associated with electronic conduction (leakage currents) through the device. As the frequency is increased, a different loss mechanism starts to dominate yielding a frequency-dependent conductance that decreases with applied field. Over most of the frequency range this second loss mechanism, which we associate with AC losses within the BST material, varies almost linearly with frequency. Thus as a starting point for a simple model we can write

$$G = G_{DC}(V) + G_{AC}(\omega, V)$$

(4.15)

The leakage term $G_{DC}(V)$ can be easily determined by a simple DC current-voltage measurement using a picoammeter. Although DC leakage is linked with long-term reliability issues, it does not usually impact the RF loss in the device under normal operating conditions. In addition, accurate modeling of leakage requires a detailed treatment of several different conduction mechanisms and the results depend critically on the choice of electrodes, process conditions, and other factors. For these reasons we will not consider leakage here and refer the reader instead to [76]-[78] and references therein for an excellent treatment of leakage in
thin-film BST capacitors. It is important to note, however, that any bias-dependent AC measurements can be influenced by leakage in the device at high fields.

From (4.14) we can see that the AC loss term, modeled as a conductance with a nearly linear frequency dependence, is consistent with a constant loss tangent over several orders of magnitude in frequency. This behavior is observed in many materials and is sometimes called “universal relaxation” [79] in which the complex permittivity follows a power-law

\[ \varepsilon(\omega) - \varepsilon_\infty \propto (j\omega)^{n-1} \]  

(4.16)

where \( \varepsilon_\infty \) is a high-frequency asymptote, and \( n \) is an exponent that is usually close to unity and can be fundamentally linked to measurements of time-dependent depolarization currents [80,81] (in the time-domain, universal relaxation is called Curie-von Schweidler behavior). Separating out the real and imaginary parts of the complex permittivity and converting to capacitance and Q-factor gives [82]

\[
C(f) = C_\infty + C_0 \left( \frac{f}{f_0} \right)^{n-1}
\]

\[
Q(f) = \frac{C(f)}{C(f) - C_\infty} \tan \left( \frac{n\pi}{2} \right) \approx \tan \left( \frac{n\pi}{2} \right) = Q_f
\]

(4.17)

where \( f_0 \) is some suitably-chosen reference frequency (usually 1 Hz) such that the fitting parameter \( C_0 \) has the units of capacitance. Thus the Q-factor in this intermediate frequency range is intimately linked with the observed dispersion in the capacitance-vs. frequency curves, as required by causality and the Kramers-Kronig relations [83].

After curve fitting the capacitance data to determine the exponent \( n \), the Q-factor or loss tangent can be determined. Typical exponents for BST capacitors are in the range of \( 0.990 < n < 0.998 \), so the change in capacitance is often slight, requiring measurements over a wide range of frequencies in order to determine accurately. Since capacitance measurements are generally more accurate than loss measurements for low-loss materials, the relationship in (4.17) can be exploited to characterize the loss and identify possible

Bob York
extrinsic contributions, or simply to establish confidence in a direct measurement of the loss. Figure 4-16a shows a typical result of such a calculation, with error bars based on the uncertainty in the fitting parameters. Similarly, Figure 4-16b compares the calculated Q-factor at 1MHz versus measured Q-factor using this technique, for materials of different thickness (from 72-380 nm), at different temperatures (from 150K-325K), and different deposition parameters (varying Ar/O₂ ratio in sputtering). The calculated Q-factors are usually slightly higher than the measured Q, suggesting some small extrinsic contribution, but overall there is excellent agreement for a wide range of films and temperatures.

In the spirit of simple modeling, we can take the capacitance and film Q-factor \( Q_f \) to be roughly constant in the MHz range, so

\[
G_{ac}(\omega, V) \approx \omega C(V) / Q_f(V)
\]

and all that remains is to characterize the voltage-dependence of the film loss, \( Q_f(V) \). Since device leakage makes such characterization difficult at low-frequencies, this will be examined in the section on high-frequency measurements.

### High-Frequency Q Roll-off

At the upper edge of the data presented in Figure 4-6b the Q-factor begins to roll off. In most cases the roll-off tends to asymptotically approach a \( 1 / f \) dependence. There are at least three possible explanations for this: 1) a frequency-dependent loss tangent, arising from some high-frequency relaxation processes in the material; 2) a series resistance due to electrodes or other extrinsic effects; 3) measurement errors due to limitations of the instrument (impedance analyzer) or calibration technique at the upper end of its measurement range. In most cases there is likely to be some combination of all three effects going on simultaneously.

It is important to understand the difficulty in distinguishing between the first two effects without additional experimentation. The parallel CG circuit in Figure 4-17 can be represented as a series RC equivalent circuit as shown, where

\[
R' \approx \frac{G}{\omega^2 C^2} \quad \text{and} \quad C' \approx C \quad \text{when} \quad \tan \delta \ll 1
\]

Remembering that the material losses can be represented as \( G = \omega C \tan \delta \), we can see that a loss tangent with a linear frequency-dependence can be modeled by a constant series resistance, which in turn leads to a Q-factor that rolls off as \( 1 / f \) since for the series circuit we have

\[
Q = \frac{1}{\omega R'C'}
\]

and thus a frequency-dependent loss tangent is indistinguishable from a series resistance arising from the metal electrodes. Note that a linear (or close to linear) frequency dependence for the loss is predicted by many simple relaxation processes such as the Debye law [83]

\[
\varepsilon(\omega) = \varepsilon_b + \frac{\Delta \varepsilon}{1 + j \omega \tau} \quad \omega \tau \ll 1 \quad \Rightarrow \quad \tan \delta \approx \omega \tau \frac{\Delta \varepsilon}{\varepsilon_b}
\]
Frequency-dependent loss tangents should be accompanied by a corresponding change in the frequency-dependence of the capacitance. This is not usually observed in thin-film capacitors; the capacitance generally appears to follow the weak power law predicted by universal relaxation in (4.17) up into the GHz range [81]. However, tunable materials are somewhat unique in that they have very large dielectric constants, such that the relaxation process contributing to the loss may have a negligible impact on the overall capacitance (equivalent to assuming that $\Delta \varepsilon / \varepsilon_b \ll 1$ in (4.21)). In fact, the small series inductance in a real device is often sufficient to mask the additional capacitive dispersion that would accompany a frequency-dependent loss tangent.

Regardless of the physical origin, it seems that we can always model the high-frequency roll-off reasonably well by adding a series resistance to the model, and then the question becomes how this term scales with geometry and applied field. The geometrical dependence of ohmic losses are relatively easy to measure and quantify for a given device structure; for example, using a distributed-circuit model [56] the series resistance for the simple test structure shown in Figure 4-5 can be written as

$$R_{\text{electrode}} \approx \frac{r_b}{2L} \left( g + \frac{W}{6} \right) + \frac{1}{3} \frac{L}{W} \left( r_t + \frac{r_t \parallel r_s}{2} \right)$$

(4.22)

where $r_b$ is the sheet resistance of the bottom electrode, and $r_t$ is the sheet resistance of the top electrode. In most cases the first term is dominant, since the bottom electrode is usually a thin refractory metal. It is relatively easy to characterize the sheet resistance of each metal layer using suitably designed process monitors, but the actual device resistance can also be estimated more directly by fabricated short-circuited devices (no dielectric) alongside the test structures. This will not include some of the distributed effects accounted for in (4.22), but these are often second-order effects.

Frequency-dependent material losses, on the other hand, should scale with area in the same way as a contact resistance term,

$$R_{\text{material}} \approx \frac{r_c(V)}{A}$$

(4.23)

where $r_c$ involves the loss tangent, has the units of specific contact resistivity ($\Omega \cdot \mu\text{m}^2$), and is voltage dependent. (A similar expression would arise from interfacial contributions to the series resistance, like that discussed in [61]). The total series resistance is then a sum of $R_{\text{electrode}} + R_{\text{material}}$. The different geometrical dependencies of each term can be helpful in identifying sources of loss.

**Other Geometrical Effects on Tunability and Q-factor**

It is important to appreciate that until this point, all the devices and results that have been considered were based on relatively large-area capacitors. This was intentional, because there are some important geometry effects that influence the electrical properties of the device as the capacitor area shrinks. Dielectric
varactors have a high capacitance density, up to 100 times that of conventional integrated capacitors using SiO₂ or SiN dielectrics. The typical electrode areas are therefore much smaller by comparison, and the periphery-to-area ratios are much higher for a given total capacitance. This is especially true for tunable RF applications because relatively small devices are required for circuit designs in the GHz range.

Experimentally we observe that smaller capacitors have a reduced tunability compared with larger devices on the same wafer. This appears to be well modeled by a non-tunable peripheral or “fringing” capacitance $C_f$ in parallel with the tunable device, as was shown in Figure 4-10. As the device size is reduced this contribution represents an increasing fraction of the overall capacitance, and the tuning curves are observed to level off prematurely. Our data is consistent with a modified $C(V)$ relationship of the form

$$C(V) = \frac{C_{\text{max}} - C_f}{2 \cosh \left( \frac{2}{3} \sinh^{-1} \left( \frac{2V}{V_2} \right) \right) - 1} + C_f$$  \hspace{1cm} (4.24)

The fringing capacitance scales with periphery and seems to have a weak thickness dependence. Figure 4-18 shows the tuning curves for several small-area devices and a comparison to the theoretical model with and without the fringing correction [57]. The devices were made using sputtered 30/70 BST with Pt electrodes. The dashed curves in Figure 4-18 were generated from (4.6) using $V_2 = 13$ Volts, a value determined experimentally from larger area devices on the same wafer. The solid curves were generated from (4.24) using a 2.8fF/μm fringing capacitance density. The data is shown on a log-scale in capacitance for clarity (note that the data was rounded to the nearest 0.01pF, which is apparent in the data for smallest device).

![Figure 4-19](image-url) - Numerical contour plot of permittivity in the region of the top contact edge for a BST varactors biased at $V = V_2$. A 2D PDE solver was used.

The fact that there is a fringing capacitance is not too surprising, but simple electrostatic modeling of the device suggests that an ordinary fringing contribution should have a negligible impact on the tuning curves, in contrast to what is observed experimentally. Figure 4-19 shows the results from a 2D numerical model of a dielectric varactor biased at $V = V_2$, assuming a zero-bias permittivity of $\varepsilon_0 = 250$; here we are showing the spatial
variation of permittivity within the material, based on a computation of the electrostatic field that included the material nonlinearity described by (4.9). Near the contact edge we can see the permittivity changes, from its intermediate value of 125 in the active region beneath the top contact, to its unbiased value of 250 in the shelf away from the top contact edge. The high fields near the top contact edge reduce the permittivity even further. According to this analysis, the fringing capacitance should tune with bias voltage and should have a negligible impact on the tunability curve, other than to increase the total capacitance accordingly. Thus other explanations must be sought for the physical origin of this fringing term.

Figure 4-20 – (a) Experimental Q-factors for different device sizes [84], and (b) plot of the same data at 1MHz versus periphery-to-area ratio.

Another geometrical effect that has been observed experimentally is an apparent area-dependent Q-factor in the MHz region [84], as shown in Figure 4-20a. For an ideal parallel-plate capacitor there can never be an area-dependence for material-related losses, because the area-dependence always cancels out in the expression for Q-factor. Assuming this is not a measurement artifact, the only way to explain the data is by introducing a periphery-dependent loss term, much like we just did for the fringing capacitance. This is shown in Figure 4-21. Adding a conductance of the form $G_p = \sigma' P$, where $\sigma'$ is an effective conductivity and $P$ is the device periphery, we can show that the inverse of the total device Q-factor for this circuit is

$$\frac{1}{Q} = \frac{G_b + G_p}{\omega C_b} = \tan \delta + \left( \frac{\sigma'}{\omega c_d} \right) \frac{P}{A}$$

where $c_d$ is the capacitance density of the film. Hence a plot of the inverse Q versus periphery-to-area ratio should yield a straight line, with an intercept that gives the bulk loss tangent, and the slope gives information on the conductivity of the periphery. Figure 4-20b is a plot of the inverse Q data from Figure 4-20a taken at 1MHz, showing excellent agreement with this simple phenomenological model. The physical origin of this peripheral loss term is not clear, nor is it important from the standpoint of modeling the effect, but it may be linked with surface conduction on the BST shelf around the top-electrode [85].

Bob York
4.3 High-Frequency Measurements and Model

**RF Device Structure and Network Analyzer Measurements**

Characterization of BST varactors at radio frequencies is complicated by several factors. At these frequencies, series inductance associated with the electrode geometry introduces a self-resonant frequency that limits the useful measurement and operating bandwidth. The high-capacitance density of the films also means that capacitors intended for use in this frequency range (typically on the order of 0.1-10pF) will have rather small electrode areas, often smaller than the tips of on-wafer probes. This has at least three important ramifications: first, an additional dielectric cross-over layer is required to make an external connection to the device; secondly, the bottom electrode must be patterned to allow for the external connections; and third, the device must be integrated into a structure with large electrical contacts, introducing additional parasitics that must be de-embedded from the measurements. Thus the fabrication of RF test structures and the measurement procedures are always more difficult in the GHz range. Further complicating the issue is the intrinsic limitation of network analyzer measurements, a notoriously inaccurate technique for highly reactive devices (impedances on the outer rim of the Smith chart). Measurement interpretation is also more difficult, as new loss mechanisms become significant in the GHz range that may not be apparent at lower frequencies (dielectric relaxation processes, interfacial losses arising from electron transfer between the electrodes and surface states, and skin-effect losses in the electrodes).

![Diagram of RF device](image)

Figure 4-22 – Simple RF device layout and photo of a finished structure. The small contact area requires a small access finger that must cross over the bottom electrode, necessitating the use of an air-bridge or dielectric cross-over layer. A U-shaped “collector” is used to minimize resistance associated with the bottom electrode [56,110].

A basic RF device structure that has been used in most of our work is shown in Figure 4-22 [56,86,87]. For the most part this is a standard integrated-circuit capacitor structure [27,28] where the top contact to the capacitor dielectric (BST) is defined by a window in a second (usually low-κ) dielectric. This second “interlayer” dielectric serves two important roles: first, as a cross-over layer to separate the top contact away from the edge of the bottom electrode, a problem region for premature breakdown; secondly, as an environmental encapsulant to protect the BST film from exposure to subsequent contamination in processing or operation. In the latter role the layer is also sometimes referred to an a “passivation” layer,
suggesting that it may be helpful in controlling surface states in the film, but this effect has not been widely studied yet. Several materials have proven adequate for this encapsulant, such as SiO_2, SiN [27,86] and Al_2O_3 [88]. Some process-related considerations are discussed in [56,86,87]. The bottom electrode is patterned (by lift-off, dry-etching, or ion-milling techniques), and is often a thin, refractory metal such as Platinum. The only slightly unusual aspect of this device design is the “U”-shaped connection to the thick metal interconnect layer, which is designed to minimize the series resistance associated with the bottom electrode layer [56,110]. This is similar to high-speed Schottky diode layouts [89].

![Figure 4-23](image)

Figure 4-23 – Simple CPW test structures and equivalent circuits for de-embedding device characteristics from one-port network analyzer measurements.

For on-wafer characterization the device must be embedded in a “probe-friendly” structure; Figure 4-23 illustrates a simple and commonly-used scheme for one-port RF device characterization using GSG probes [90,91]. In order to remove the influence of the large CPW probe structure on the device measurement, two additional structures are fabricated alongside the device-under-test (DUT) in which the DUT is replaced by an open- or a short-circuit. If the probe pads are represented electrically by the L-network of \( Z_s \) and \( Y_p \), then the short-circuit measurement yields \( Z_s \), and the open-circuit measurement then yields \( Y_p \). Once these are known, their contribution to the measured impedance of the DUT can be mathematically removed. It has been shown in [91] that inconsistent placement of the probe tips between the various test structures can lead to measurement errors; these can be compensated to some extent, but a simple safeguard is to include some reference marks in the test structure as shown in Figure 4-23.

Usually the smallest probe-pitch (separation between the probe tips) is used for optimal accuracy. Somewhat simpler test structures are made possible if two-conductor GS (ground-signal) probes are used instead of GSG probes [92], but accurate calibration is a more difficult challenge with GS probes and will not be considered here. For one-port network analyzer measurements, a standard short-open-load (SOL) technique is typically used, with calibration “standards” provided by the probe manufacturer. Unfortunately this technique is highly sensitive to the accuracy of the impedance standards [93], especially for highly reactive and low-loss devices (reflection coefficients close to unity). Furthermore, even with perfect standards, the intrinsic measurement accuracy of the network analyzer is also suspect for such devices [94]. Two approaches that can improve measurement accuracy are the use of additional on-wafer lumped-element structures [95] as part of an over-determined...
calibration procedure, and the use of two-port structures [96] as shown in Figure 4-24. Two port methods help in two ways: first, they allow for more advanced calibration methods such as the Through-Short-Delay (TSD) and Line-Reflect-Match (LRM) techniques, which do not require perfectly characterized standards; and secondly by bringing the impedance trajectory in closer to the center of the Smith chart (closer to 50 Ohms) where the measurement accuracy of the network analyzer is significantly higher.

In a two-port measurement the DUT can be embedded in series or in shunt. The CPW shunt arrangement has the disadvantage of requiring two devices which must be assumed identical, but has the advantage of simpler biasing than the series configuration. The series configuration tends to have better accuracy when the DUT impedance is low, so it tends to work well at higher frequencies for capacitive devices; conversely the shunt configuration has better accuracy at the low end of the band when the device impedance is high. In practice, measurements taken from a combination of series and shunt structures will yield the best accuracy over a wide bandwidth.

**Curve-Fitting to a Circuit Model**

Although swept-frequency network analyzer measurements are not as accurate as some other techniques (notably high-Q resonator methods), they provide information on the behavior of the devices over a broad frequency range, which is invaluable for developing models and exploring various contributions to the device impedance.

Based on our low-frequency measurements and physical expectations, dielectric varactors should be reasonably well-modeled by the equivalent circuit shown in Figure 4-25. The only new addition to the model from the previous section is the series inductance, which can no longer be neglected in the GHz range. To estimate these model parameters we first compute the intrinsic device impedance by de-embedding the probe parasitics at each frequency (using either the one-port or two-port techniques described earlier), and then use a least-squares curve-fitting procedure following [97]. The equivalent circuit leads to an expression for the frequency-dependence of the reflection coefficient.
\[
\Gamma(\omega) = \frac{1 + (F_1 - Z_0 F_0)\omega + F_0 \omega^2}{1 + (F_1 + Z_0 F_0)\omega + F_0 \omega^2}
\] (4.26)

where

\[F_0 = jC \quad F_1 = jRC \quad F_2 = -LC\]

and \(Z_0\) is the reference impedance (e.g., 50Ω). Note that the loss tangent of the material is accounted for by allowing the unknown capacitance to be complex, \(C(1 - j\tan\delta)\). If measurements on each standard are made at \(N\) frequencies \(\omega_i\) yielding a de-embedded reflection coefficient \(\Gamma(\omega_i)\), then (4.26) leads to the following over-determined system

\[
\begin{bmatrix}
A_1 & B_1 & C_1 \\
A_2 & B_2 & C_2 \\
\vdots & \vdots & \vdots \\
A_N & B_N & C_N
\end{bmatrix}
\begin{bmatrix}
F_0 \\
F_1 \\
F_2
\end{bmatrix}
= 
\begin{bmatrix}
1 - \Gamma(\omega_1) \\
1 - \Gamma(\omega_2) \\
\vdots \\
1 - \Gamma(\omega_N)
\end{bmatrix}
\] (4.27)

where

\[A_i = \omega_i Z_0 [1 + \Gamma(\omega_i)] \quad B_i = \omega_i [1 - \Gamma(\omega_i)] \quad C_i = \omega_i^2 [1 - \Gamma(\omega_i)]\]

This equation can be solved for a least-squares best fit for the three unknowns \(F_0\), \(F_1\), and \(F_2\), and hence the model parameters from (4.26). Typically a singular-value decomposition (SVD) algorithm [98] is used. A wide frequency range and large number of measurement points are preferred for a good model fit. Note that a frequency-dependent loss mechanism (such as skin-effect losses) can also easily be included in this model if desired.

![Figure 4-26 – Example curve fit to network analyzer data from 200MHz-40GHz using the equivalent circuit of Figure 4-25 [56].](image)

Figure 4-26 shows the results of this procedure using broadband on-wafer RF data measured from 50MHz to 40GHz on an Agilent E8364A PNA-series network analyzer [56]. After a standard on-wafer SOL calibration, the device impedance was determined by first de-embedding the probe pad parasitics, and then the fitting procedure was applied, with the best-fit model parameters shown. Excellent agreement is observed between the data and the simple model of Figure 4-25. The influence of the small series inductance is clearly visible. Using the simple model in Figure 4-25, all of the high-frequency rolloff in Q-factor is attributed to the series resistance and inductance, such that the parallel-conductance term is represented by a constant loss tangent.
RF Loss Modeling and Electrostrictive Resonances

At the low end of the RF frequency range the Q-factors asymptotically approach loss tangents that are consistent with observations and models developed at lower frequencies, taken as a frequency-independent constant \( Q_f = 1/\tan \delta \) in the model. As we showed earlier in connection with (4.19), a frequency-dependence in the loss tangent can effectively be accounted for in the series resistance term, \( R_s \). This combination of loss terms in the model seems to adequately account for the observed losses in the low GHz range.

The model parameter that immediately jumps out of the example in Figure 4-26 is the excessive series resistance \( R_s \). This number is significantly larger than what would be expected from electrode contributions alone, suggesting another dominant loss term. One simple approach to identifying possible loss mechanisms is to explore the dependence on various geometrical factors. For example, we expect certain losses to scale with aspect ratio (e.g. top electrode resistance), or periphery (e.g. bottom electrode and/or fringing effects), or area (e.g. material or interface effects). We might then assume a series resistance of the form

\[
R_s \approx r_c \ell + \frac{r_c}{w} + \frac{r_c}{P} + \frac{r_c}{A}
\]

(4.28)

where the coefficients for each term are determined by experiment. Figure 4-27 shows the net series resistance determined by model-fitting of broadband RF data from 200MHz to 5GHz, using a large matrix of 144 devices with rectangular contacts varying in both length and width from 5\( \mu \)m to 40\( \mu \)m, and aspect ratios \((L/W)\) varying from 40/5 to 5/40. For this particular material and process technology there seems to be an inverse area dependence for these devices, which would suggest either a frequency-dependent loss tangent or interfacial contribution that is large in comparison to other terms. This is consistent with observations of other researchers [99]-[103]. We will not speculate further on the origin of the loss, but simply note that the zero-bias RF loss seems to be well modeled by a term of the form \( R_s \propto \frac{1}{A} \), where the effective contact resistance appears to depend on several technological variables such as grain size and/or defect concentration. With careful optimization of the material and device structure this term can be reduced. Relatively high RF Q-factors have been demonstrated from well-designed devices [104].

Under bias, both the low-frequency asymptotic loss tangent and the effective series resistance vary with applied field as shown in Figure 4-28. The variation in series resistance (which further argues for a material or interface contribution) is relatively weak and can be ignored without too much error, or fitted to a linear dependence if desired. The background loss tangent seems to obey a parabolic model reasonably well such that
where \( Q_0 \) is the zero-bias asymptotic Q-factor, and \( V_q \) is a fitting parameter (note we could replace voltage by field if desired). \( Q_0 \) is essentially the same as that determined by low-frequency impedance analyzer measurements, and may display similar dependencies on film thickness and device geometry.

\[
Q_f(V) = Q_0 \left[ 1 + \left( \frac{V}{V_q} \right)^2 \right]
\]

(4.29)

Figure 4-28 – Variation of (a) the asymptotic quality-factor \( Q_f \), and (b) the effective series resistance \( R_s \) with the applied field or bias.

At high fields the data and (4.29) diverge; it is difficult to ascribe any significance to this observation because of the inherent inaccuracy in measuring such high Q-factors on a network analyzer, and also because of the possible influence of leakage on the data at extremely high fields (see previous section). Fortunately this is a non-issue because as \( Q_f \) becomes large, its impact on the model predictions becomes small and the series resistance term is dominant. So (4.29) works reasonably well in a practical sense.

Figure 4-29 – Voltage-induced resonances in two different BST capacitors: (a) a 30×30\( \mu \)m device with 100nm Pt top electrode. (b) 10×30\( \mu \)m device with 100nm/1\( \mu \)m Pt/Au top electrode.

Most importantly, a new effect appears under bias that was not apparent at low-frequencies: resonant dips that appear in the reflection coefficient data, with a depth that varies with bias, and at frequencies that depend on the device layer structure. Examples from two devices are shown in Figure 4-29. This is now understood to be a consequence of the electrostrictive property of the material, which has been known for some time [1,105] but seemed to have escaped notice in thin-film research community until recently [106,107]. The
extent to which this effect is observable (note the scale in Figure 4-29) depends on the film thickness, electrode area, and applied field.

This electrostrictive property is akin to a piezoelectric effect, but depends on the applied field so it can be called “field-induced piezoelectricity”. The piezoelectric coupling coefficients for the material can be quite large and seem to vary approximately linearly with applied field. If the piezoelectric coupling constants are known, the BST varactor can be modeled as a thickness-mode bulk-acoustic-wave (BAW) device using standard models such as the KLM model \[108\] or the Mason model \[109\]. The KLM model is attractive because each of the various layers in the vertical cross section of the device is modeled as an equivalent transmission-line. Figure 4-30 shows a comparison of the modeled reflection coefficient for a typical varactor structure using three different top electrodes.

![Diagram of varactor structure](image)

Figure 4-30 – Measured reflection coefficient of a 31×32 μm device, 160nm BST, at 20V bias, for three different top electrode stacks, illustrating the damping effect of a thick gold layer on the electrostrictive resonances (courtesy of Agile Materials & Technologies).

Although this property of STO and BST is quite interesting and potentially quite useful for high-Q resonators, we will not consider it further in this work since our focus is on tunable capacitors. As Figure 4-30 indicates, the resonances can be effectively suppressed to a large extent by piling on thick layers of interconnect metal, which acts like a mechanical damper on the structure. There is still an adverse residual effect on the loss properties of the device, but this can effectively be accounted for by existing terms in our simple loss model.

**Design Optimization**

With the various dependencies of loss and tunability on geometrical variables articulated, we can begin to focus on designs that optimize certain performance variables.

Our analysis already began with a device structure that minimizes electrode resistance (Figure 4-22) with a thick metal collector surrounding the top contact. As a result, the net Q-factor for the device was dominated by mostly area-independent effects, so we can not expect further dramatic improvements in Q-factor with device design, except perhaps by decreasing the sheet resistance of the electrode layers \[104\]. Focusing then on secondary effects, two small contributors to the high-frequency roll-off are the top electrode resistance, which includes the “access” resistance associated with the length $\ell_a$ in Figure 4-22, and the series inductance $L_s$, each of which scale with the aspect ratio $L/W$; this tends to argue against long, narrow top contacts. The inductance contributes to high-frequency roll-off through the self-resonant frequency, given by
The self-resonant frequency is given by

$$\omega_r \approx \frac{1}{\sqrt{L_s C}}$$

Data indicates that the series inductance is roughly area-independent, so maximizing the self-resonant frequency requires the smallest possible capacitance. Small-area capacitors are therefore expected to perform best with respect to high-frequency roll-off in $Q$, and this is indeed borne out by experiment. On the other hand, we have also seen that the constant background loss tangent (associated with $Q_Q$) may have an area-dependence such that it increases with area-to-perimeter ratio. In addition, the tunability also increases with area-to-perimeter ratio because of the non-tunable fringing capacitance; this effect is shown in Figure 4-31.

A partial solution to these competing demands is to use a number of small-area devices in parallel to achieve a given capacitance, and to use a top-electrode geometry for each unit cell that maximizes the area-to-perimeter ratio for a given area; this argues for circular top electrodes. Circular electrodes have also shown to be better than rectangular electrodes from the standpoint for breakdown and long-term reliability in thin-film capacitors.

Another reason for using multiple parallel devices is to minimize the current density in the electrode layers [110]. This can be important for applications requiring large RF voltage swings. It is well known from studies in integrated circuit technology that the current density must be kept below some critical value $J_c$ to prevent failure

$$J < J_c$$

The critical current density depends on the type of metal and its thermal environment, but is typically on the order of $10^6$ A/cm$^2$. The amount of AC current that will flow in the capacitor is a function of the RF voltage and impedance. If the peak AC voltage swing is denoted by $V_{\text{max}}$, the peak AC current through the capacitor is

$$I_{\text{max}} = j\omega CV_{\text{max}}$$
For a single device with a round top electrode of radius $r$, top interconnect thickness $t_c$ and bottom electrode thickness $t_b$, then (4.31) implies the following approximate inequalities for the top and bottom metal layers

$$\frac{I_{\text{max}}}{t_c 2r} < J_c$$

$$\frac{I_{\text{max}}}{t_b \pi r} < J_c$$

(4.33)

Since we usually have $t_b \ll t_c$, it is the second inequality that is more restrictive. Using (4.32) and writing $C = \varepsilon_r \pi r^2$ we find

$$r < \frac{J_{\text{tr}} t_b}{\varepsilon_r \pi V_{\text{max}}}$$

(4.34)

As an example, consider an application at 1GHz with an 10Vp-p signal swing (corresponding to 1Watt in a 50Ω system). If we assume a bottom electrode thickness of $t_b = 300\text{nm}$, a critical current of $10^6 \text{A/cm}^2$, a capacitance density of $c_d = 20 \text{fF/μm}^2$ gives $r < 2.4 \mu\text{m}$, corresponding to a capacitance of about 0.4pF. If the application requires a capacitance of 2pF, this would have to be realized using parallel combinations of small-area devices.

**Linearity and Control Voltage Tradeoffs**

The current-handling analysis leading to (4.34) indicates that a low capacitance density is desirable. We similarly found (Figure 4-14) that films of in excess of 100-200nm are also desired to maximize tunability for a given maximum field strength. Low capacitance density also helps from a processing standpoint, especially for high frequencies where the required capacitance values are small and the electrode areas become small. The downside of thicker films, however, is a larger DC control voltage. Some applications, especially in wireless handheld devices, would prefer to have low control voltages commensurate with the battery technology, which may be as low as 2.7V.

The tradeoff with control voltage is especially important with respect to linearity concerns in the RF front-end. The capacitance change with voltage occurs almost instantly in tunable dielectrics, so a large RF voltage swing across the device will modulate the capacitance and generate harmonic distortion, including third-order products (inter-modulation distortion, or IMD) if the waveform is modulated. IMD is a particular concern in modern wireless systems with close channel separation such as GSM and CDMA. Devices with low control voltage will generally produce more harmonic distortion because, by definition, they generate a larger capacitance change per volt. **So the desire for low control voltage is always at odds with the desire for high power handling and high linearity.**

One solution to this problem is to use a simple DC-DC boost-converter or charge-pump to step-up the available control voltage to a much higher value, large enough to drive a dielectric varactor with sufficient power-handling for the application. This is in fact a reasonably attractive option for some applications, because dielectric varactors draw no appreciable current and hence consume almost no power from the control circuit (only during switching transients is there any significant current draw). DC-DC converters can be quite small and inexpensive when they do not have to source much current, and can be designed for quite large step-up ratios.

However, the linearity specifications for many applications simply require film thicknesses and control voltages that are too large to be practical, and hence a different solution is needed. The best solution to this problem appears to be the use of a series combination of capacitors, as shown in Figure 4-33 [111]. In this figure, the resistors are all large enough that they are effectively not part of the RF circuit, so the total RF voltage is
divided more-or-less equally among each of the capacitors. However, the biasing circuit is designed such that the *full* applied DC control voltage appears across each capacitor. There are a number of possibilities for accomplishing this, depending on whether the varactor is mounted in series or shunt in the RF circuit, and depending on whether the bias voltage can be superimposed on the RF signal. Other straightforward extensions of this concept (not shown) might include a large-value DC blocking capacitor to isolate the bias circuit from the RF circuit, or perhaps using inductive chokes instead of resistors. Vertical stacking has been demonstrated with both interdigital and parallel-plate capacitors [112-113].

![Figure 4-33](image1.png)  
**Figure 4-33** – Examples of vertical device stacking to enhance power handling and linearity with low control voltage [111]. (a) Two-terminal structure with bias and RF superimposed at one terminal. (b) three-terminal arrangement with separate bias and RF circuits.

![Figure 4-34](image2.png)  
**Figure 4-34** – A high-linearity varactor design using stacking and parallelism for high power handling and low control voltage (photo courtesy of Agile Materials & Technologies Inc.)

The vertical stacking concept can be scaled to arbitrarily large numbers of devices in series, so there is great flexibility in manipulating linearity and power-handling with respect to control voltage. Note that if there are a total of *N* capacitors *C_s* in series, then the net capacitance of the structure is reduced to *C_s/N*, so as the stacking increases, the unit cell capacitance must also increase. In order to maximize the Q-factor and current-handling, it is advantageous to combine the parallel device concept described in Figure 4-32 with the series stacking described by Figure 4-33. An example of such a device is shown in Figure 4-34.
There are at least two practical challenges with the use of stacked capacitors; first, the need for a number of large value resistors, which can occupy significant substrate area unless a high sheet-resistance process is available. Secondly, as the number of devices grows, the settling time with respect to a step-function control voltage increases, because the inner devices in the stack must be charged through an increasing number of resistors. This tends to favor smaller resistors, but small resistors decrease the overall Q-factor and increase the leakage currents, so there is an inherent tradeoff between Q-factor and settling time. This can be partially addressed using alternative biasing schemes that will not be discussed here.

**Q-Enhancement**

Tunability and material Q-factor are fundamentally linked, such that the choice of material always involves a compromise between the two factors. But occasionally it is more effective to use circuit techniques to manipulate this tradeoff. An example is a series combination of high-Q non-tunable capacitor with a low-Q tunable capacitor. Many thin-film processes include some kind of low-κ dielectric layer that can be used for a high-Q fixed capacitor, making this a very practical method for enhancing the overall Q-factor at the expense of some tunability.

Let's represent the high-Q non-tunable capacitor as \( C_q \) with a Q-factor \( q \), in series with a BST capacitor \( C_d \) of Q-factor \( d \). If the BST capacitor has a tunability of \( \tau \) over some control voltage range, then we can show that the overall Q-factor and tunability are of the series combination is given by

\[
Q = \frac{Q_d Q_q (C_q + C_d)}{C_q Q_q + C_d Q_d}
\]

\[
\tau = \frac{C_d + \tau C_q}{C_d + C_q}
\]

(4.35)

The effective Q-factor versus tunability for the series combination is shown in Figure 4-35, assuming a BST tunability of 3:1, a BST Q of 100, a non-tunable device Q of 500, and an overall net capacitance of 1pF. We can see that this technique can be effective for Q-enhancement in situations where Q is at a higher premium than tunability. It should be noted that similar ideas have been used to enhance Q using layered materials rather than distinct capacitor structures [64]; recently, Yan et al. [114] have also demonstrated this concept with a layered structure of BZN and BST dielectrics.

**4.4 Phase-Shifters and Delay Lines**

Early interest in thin-film BST focused on the potential use of this technology in phase-shifters for low-cost phased-array antennas. Compared with conventional MMIC processing, BST varactor circuits have a potential for fewer processing steps, cheaper substrates, and larger lithographic dimensions, all factors that reduce costs. In addition, BST-based phase-
shifters provide continuous analog phase control with a single control voltage, negligible control power dissipation, and fast response times. Such analog functionality is especially useful in beam-scanning applications for improving beam-pointing accuracy and compensating for temperature or aging drift. BST-based phase-shifters circuits do not require hermetic packaging and are compatible with flip-chip manufacturing techniques. Collectively these features make BST phase-shifters an attractive alternative to competing MMIC or MEMs technologies.

Key requirements for a viable BST phase-shifter technology are low insertion loss, wide bandwidth, and small die size. This section compares some possible design approaches with respect to these criteria.

**Periodically-Loaded Transmission-Lines**

Traveling-wave phase-shifters were among the first types of structures to be examined using BST materials [115]-[122]. Here a transmission-line is loaded with BST material in such a way that the phase-velocity is dependent on the permittivity of the material, and hence the propagation delay along the structure is controlled by the DC applied field to the material. Early circuits typically used a CPW or microstrip arrangement, where the BST material continuously loads the transmission line. This creates a difficult impedance matching problem because of the high permittivity of the material.

![Figure 4-36 – Periodically-loaded transmission-line structure](image)

A more attractive solution is shown in Figure 4-36, where discrete BST varactors periodically load a transmission line (CPW shown). This was first demonstrated using Schottky diode varactors [122]. A design optimization procedure is described in [122] and summarized here. It is helpful to define a “loading factor” $x$, which is the ratio of the loading (varactor) capacitance to the unit-cell capacitance of the unloaded line

$$x = \frac{C_d(V)}{C_{cell}}$$  \hspace{1cm} (4.36)

The impedance and phase velocity of this structure are then given by

$$Z_\perp = \frac{Z_0}{\sqrt{1+x}} \hspace{1cm} \nu_p = \nu_0 \sqrt{1+x}$$  \hspace{1cm} (4.37)
where $Z_i$ and $v_i$ are the impedance and velocity of the unloaded transmission line, which can be related to $L'$ and $C'$, the inductance and capacitance per unit length as

$$v_i = \sqrt{L'C'} \quad Z_i = \frac{L'}{\sqrt{C'}} \quad (4.38)$$

The periodic nature of the structure introduces a cutoff frequency called the Bragg frequency,

$$f_{\text{Bragg}} = \frac{v_i}{\pi \ell_{\text{cell}} \sqrt{1 + x}} \quad (4.39)$$

In writing (4.37) we basically assumed that the loading capacitance is spread uniformly over the unit cell, so these expressions only work well below the Bragg frequency. In this case, the maximum differential phase-shift per unit cell is given by

$$\Delta \phi = 2\alpha \ell_{\text{cell}} \left( \sqrt{1 + x} - \sqrt{1 + x/\tau} \right) \quad (4.40)$$

and the total loss per unit cell can be written as

$$\alpha \ell_{\text{cell}} = \pi Z_{Lc} C_d \max \left( \frac{f}{Q_f} + \frac{f^2}{f_c^2} \right) + \alpha \ell_{\text{cell}} \frac{Z_i}{Z_L} \quad (4.41)$$

where $\alpha$ is the attenuation factor of the unloaded line, and we have defined a device cutoff frequency $f_c$ due to ohmic loss in the BST varactor as

$$f_c = \frac{1}{2\pi R_f C_d \max} \quad (4.42)$$

We can combine (4.40) and (4.42) to give a figure-of-merit (FOM) as

$$\text{FOM} = \frac{2\pi f \left( \sqrt{1 + x} - \sqrt{1 + x/\tau} \right)/8.686}{\pi x \sqrt{1 + x} \left( \frac{f}{Q_{\text{film}}} + \frac{f^2}{f_c^2} \right) + \alpha_v \sqrt{1 + x}} \quad [\text{deg/dB}] \quad (4.43)$$

and the overall length of the structure that is required to achieve a 360 degree phase delay at some frequency $f_0$ is given by

$$L_{\text{total}} = \frac{v_i / f_0}{\sqrt{1 + x} - \sqrt{1 + x/y}} \quad (4.44)$$

For a given set of design variables the FOM increases with loading factor, and the total length decreases with loading factor. The design frequency is chosen as the frequency which maximizes the FOM. In order to maintain a good impedance match, the loaded impedance must remain close to 50Ω over all bias states. For large loading factors, this requires that the unloaded line impedance be quite large. In addition, the return loss begins to increase dramatically near the Bragg frequency, so the operating frequency must be chosen well below the Bragg frequency, usually around 25-50% lower.

Early designs using coplanar waveguide [123]-[129] were limited to small loading factors because of the limited range of CPW impedances that can be realized with low loss. Some impressive FOM results (approaching 90°/dB) were achieved [128,129], but these structures were quite large. A solution to this problem is to use coplanar-stripline (CPS), the dual of CPW. CPS can be realized with quite large unloaded impedances and with relatively low attenuation, allowing for high loading factors and hence more compact designs. Figure 4-37
shows the FOM calculation and a measured result for a 12GHz CPS-based design. This device measured 1×3mm overall, considerably smaller than CPW designs. Similar structures were used recently in a 24GHz wafer-scale phased-array system [130].

For high tunability and/or high loading factors, the loaded impedance of the phase-shifter can vary significantly. A simple partial solution to this problem involves using smaller varactors in the first and last sections, effectively providing a small amount of impedance transformation. Another approach using a different unit cell design is described in [131].

**LC Ladder Structures**

The wide bandwidth and parameter-insensitive design make distributed delay-lines attractive, but there are two serious drawbacks: first, the designs are physically large, increasing the unit cost and complicating array designs; secondly, both the CPW and CPS structures require off-chip baluns for proper functioning in the microstrip environment common for a host circuit.

Figure 4-38 – Two small variable delay-line structures using an LC ladder network, and representative s-parameters at various bias states (after [135]).
One way to reduce the size of the structure is to implement the transmission-line as a synthetic LC ladder network, using lumped-element inductors for the series element. Again, this technique has been implemented using Schottky varactors [132,133] as well as BST varactors [134,135]. Figure 4-38 shows two small BST-based phase-shifter structures using this LC ladder approach. These designs yielded approximately 90° phase shift with ~2dB of loss at the design frequency of 10GHz, corresponding to an FOM of approximately 45°/dB, with excellent return loss over the bias states as shown. Similar structures were used recently in an adaptive amplifier linearization network, reported in [136,137].

The principal drawback of using spiral inductors is their increased loss compared to transmission lines, but optimization with electromagnetic field solvers can be carried out to find the inductor geometry with the lowest series resistance. Note that the optimization procedure for the overall FOM is somewhat different here than for the truly distributed structure, because the assumption of constant line loss regardless of Bragg frequency is no longer true. A lower $f_{Bragg}$ design using a smaller number of higher valued inductors may incur more loss than a design using a larger number of smaller inductors.

**Reflection Phase-Shifters**

There are many other types of phase-shifters that have been developed to exploit varactors [138]. One example is a so-called reflection phase-shifter, shown in Figure 4-39 [139,140]. In this circuit, a quadrature hybrid circuit is used with two of the ports terminated in reflecting loads. The hybrid splits the input signal, sending it to the reflecting loads, and the reflected signals then add coherently at the output port. The phase is varied by implementing the reflecting loads as variable reactance networks, some examples of which are shown.

![Reflection phase-shifter concept](image)

- Simple, limited phase shift (40° with 2:1 tuning)
- Increased phase shift, requires $\lambda/4$ line
- Increased phase shift, low-Q spiral inductor

Series LC resonators with spiral inductors are attractive for compact circuits. In this case, phase-shifters can be made with far fewer varactors than is needed for the delay-line structures described earlier. For a given varactor value $C_v$ and tunability $\tau$, it can be shown that the phase-shift and the required inductor value is given by

$$
\Delta \phi = 4 \tan^{-1} \left[ \frac{1}{2 \omega_0 Z_0 C_v \sqrt{\tau}} \left( \sqrt{\tau} - \frac{1}{\sqrt{\tau}} \right) \right] 
$$

$$
L = \frac{1}{2 \omega_0^2 C_v \tau} \left( \sqrt{\tau} + \frac{1}{\sqrt{\tau}} \right)
$$

(4.45)
where $C_{v0} = \frac{C_{\text{max}}}{\sqrt{2}}$ is the geometric average of the maximum and minimum varactor capacitance, and $Z_0$ is the characteristic impedance looking into the ports of the branchline coupler at the location of the reflecting loads. Large phase-shifts can theoretically be realized given a large enough tunability, but in practice the inductor values can become too large, so phase-shifts from 90-200° are realistic for simple resonators.

The quadrature hybrid can be implemented in a number of ways. Using a Lange coupler [141] is one common method. Figure 4-40 shows a reflection phase-shifter using a Lange coupler and BST varactors [134,142]. This particular network achieved a maximum of 120° of phase with a little over 3dB of insertion loss at around 11GHz, corresponding to a FOM of $\sim$40°/dB. The return loss was excellent over the entire 8-12GHz range. The circuit occupied a die area of 3.3mm × 0.65mm, dominated of course by the Lange coupler which must be on the order of a quarter-wavelength long. Similar designs have also been reported using inter-digital BST varactors [143].

![Figure 4-40](image)

Figure 4-40 – A reflection phase-shifter using a Lange coupler and series resonators [147].

A branchline coupler is another possibility for implementing a quadrature hybrid, but this is even larger than the Lange coupler so it is not attractive for compact monolithic circuits. An interesting alternative is to implement the hybrid using lumped-elements [144,145]. A diode varactor implementation using a lumped-element hybrid was described in [146], demonstrating that extremely compact designs could be realized with this approach. As a further refinement, note from the design equations that the phase-shift (4.45) scales inversely with the characteristic impedance of the ports; if an impedance-transforming hybrid design is used [145], the load ports can be designed for low impedance while maintaining a good impedance match at the RF input and output. This has the further advantage that smaller inductor values can be used, which in turn makes the circuit smaller and reduces some of the
spiral parasitic capacitances that conspire to reducing the phase shift. This technique was adopted in [142,147], and resulted in an extremely compact phase-shifter, shown in Figure 4-40. This device measured 0.6mm×0.6mm, and produced over 200° of phase shift.

![Image](image_url)

Figure 4-41 – A very compact reflection phase-shifter using BST varactors and an impedance-transforming lumped-element hybrid design [142,147]. This circuit measured 0.6mm × 0.6mm.

A drawback of the lumped-element hybrid implementation in comparison to the Lange coupler example of Figure 4-40 is a reduced bandwidth and higher loss (lower figure-of-merit). However, the small size may make this a more attractive approach in some applications. In both cases, the reduced number of varactors needed in comparison to the distributed delay-line architectures may be attractive from a yield and reliability standpoint.

**All-Pass Networks**

The last kind of lumped-element phase-shifter considered here is based on 2nd-order all-pass networks as shown in Figure 4-43. If we note that the LC ladder network described earlier is basically a low-pass filter, the Bragg-frequency (or cutoff frequency) is the source of some trouble, leading to a strong increase in insertion loss and limiting the amount of phase-shift that can be achieved from each unit cell. The all-pass network solves this problem by adding a bridging element to counteract this increase in loss. If designed properly, this network can give a flat amplitude response over all frequencies, with a greater phase-variation than either a simple low-pass or high-pass network. Lumped-element values that achieve this are shown in the figure for the ideal case, where $R$ is the desired impedance level (usually 50Ω) and $\omega_0$ is the center-frequency at which the maximum phase variation occurs.

\[
L_1 = \frac{R}{\omega_0}, \quad C_1 = \frac{1}{R\omega_0}, \quad L_2 = \frac{R}{2\omega_0}
\]

Figure 4-42 – Bridged-tee circuits and design equations for realizing 2nd-order all-pass transfer functions. (a) Bridged low-pass tee. (b) Bridged high pass tee.
The all-pass network has historically found use in phase equalization networks, and has also been used in varactor-based phase-shifters using diodes [148], GaAs FETs [149,150], and more recently with tunable dielectric materials [151,152]. An example of a simple bridged high-pass structure is shown in Figure 4-43. Non-idealities in the circuit, principally losses in the varactor and spiral inductors, lead to a slight dip in the insertion loss near the center frequency (around 5 GHz in this example), but the small variation in insertion loss over the bias states is an improvement compared with phase-shifters discussed earlier.

Figure 4-43 – Simple all-pass phase-shifter structure (0.25mm²) and measured characteristics (courtesy of Agile Materials and Technologies Inc.).

Like the lumped-element reflection phase-shifter this is a narrowband design, but the all-pass structure is very attractive in terms of giving a large phase-shift for a given size and varactor tunability, is extremely easy to design, and uses a minimum number of varactors.

4.5 Tunable Filters and Matching Circuits

**Tunable Bandpass Filters**

The possibility of using BST varactors in tunable filters has been recognized for quite some time, but this application space has not been as well developed or investigated as phase-shifters, although it could be argued that many of the phase-shifter structures discussed in the previous section are essentially tunable filters themselves. The reason is two-fold; first, high quality filters—that is, high-order reactive circuits with sharp pass-band skirts and large out-of-band rejection—are very difficult to design and implement, being quite sensitive to parasitic effects in the components. Secondly, most filter applications place a very high premium on insertion-loss in the pass-band, and the technology for BST varactors was simply not mature enough to satisfy the difficult requirements on Q-factor and tunability.

Nevertheless, some good progress has been made in this area. Some of the first efforts to demonstrate tunable bandpass filters with BST technology were reported by researchers active in the high-temperature superconductor field, and typically used structures involving transmission-line resonators, usually with high-Tc superconducting electrodes but with some reports of good performance at room temperature with normal conductors [153]-[157]. More recent circuit demonstrations [158]-[162] have focused on filters using integrated thin-film varactors and have continued to improve on the early results.
High-quality filters require high-Q resonator structures, and this involves some direct and indirect design challenges. First, even if the BST varactor Q is high, it must always be combined with an inductance of some kind (lumped or distributed), lowering the overall Q. For on-chip inductors such as spirals, the Q-factors are usually limited to 30-40 in the GHz range; this then becomes the dominant loss in the resonator and filter structure. Distributed (transmission-line) resonators have higher Q-factors, but not usually in excess of 100 due to a combination of ohmic loss, parasitic substrate modes, and radiation (some completely shielded resonators have Q’s approaching 200 but are large). Transmission-line structures also become prohibitively large below a few GHz. A practical solution in this frequency range is to use discrete off-chip air-core inductors with Q’s in the range of 100-200.

A second issue, sometimes overlooked, is that high-Q resonators tend to amplify the RF voltage across the varactor (or current depending on whether it is a series or parallel resonator), thus raising significant power-handling and linearity concerns. These can be addressed to some extent by the ideas developed in earlier sections, but generally involve tradeoffs with control voltage and Q-factor, among other considerations. Lastly, since only the capacitors are tuned and not the inductors, it becomes difficult to maintain a desirable passband shape and out-of-band rejection as the device is tuned, and careful optimization of the circuit is usually required.

With careful attention to detail in the design of the structures, some impressive results can be obtained [162]. Figure 4-44 shows an example of a lumped-element IF filter using air-core inductors and BST varactors. Within each separate band the frequency tuning is accomplished with an analog voltage control to the varactors. In order for the filter to operate over multiple IF bands, three different varactor sizes are used, which are connected to the RF circuit using a simple switching matrix. An advantage of off-chip inductors is that they can be mechanically adjusted to optimize the response shape after assembly. Using tapped inductor matching at the input and output of the circuit also allows for the resonator impedance to be optimized somewhat independently of the external matching considerations.
Impedance Matching Networks

The use of BST varactors in impedance-matching networks is a more recent suggestion that may prove useful in RF front-ends for optimizing amplifier and antenna efficiency.

In modern wireless systems the efficiency of the RF power-amplifier has a strong influence on battery-life. Ordinarily the amplifier is designed to operate with maximum efficiency at its maximum power level, but in a typical wireless system the amplifier is more commonly operating at a much lower power and efficiency. The solution is to use a dynamic load-line that can be adjusted to maintain high power-added efficiency at different power levels, thus keeping the current draw from the battery to a minimum.

Figure 4-45 – A simple impedance matching network operating at 900MHz. $L_2$ is an off-chip inductor, the remainder of the circuit is an integrated chip with BST varactors (1mm$^2$) [163].

Figure 4-45 is a simple impedance transforming network that illustrates the concept [163]. The circuit uses two BST varactors and an off-chip inductor. The measured data is taken using a 50Ω load, and shows that the input impedance is can be varied from 13–29Ω using a 2:1 change in the capacitance, and remains on or near the real axis over the measurement frequency range.

Figure 4-46 – An impedance-matching network for antenna tuning (chip size 0.5 mm$^2$) [163].
Figure 4-46 shows a similar impedance matching network, designed for matching to a small helical antenna in the 420-490MHz range [163]. The problem being addressed in this example is two-fold: first, it is ordinarily difficult to create an efficient match to an electrically-small antenna over a wide frequency range, due to the large reactance and small radiation resistance of the structure and fundamental Bode-Fano limitations [164]. So, using a tunable matching network allows for a less complicated and more efficient narrow-band matching network to be used, which can then be tuned to operate at various channels within the band of interest. The measured response for various combinations of tuning voltages is shown in Figure 4-46, and indicates that a good (<-10dB) match can be achieved over the range of 420-490 MHz as desired.

A second problem that is addressed with a tunable antenna matching network is the potential variation in the antenna impedance as its operating environment is changed. Figure 4-47 shows how the impedance of a simple helical antenna can vary as it is moved from the head to a table (see Smith chart in figure). The accompanying data shows how the measured return loss can vary if the impedance match is fixed, and how it can be subsequently tuned to improve the match. The ability to compensate for antenna mismatch could significantly enhance the performance of any RF front-end.

### 4.6 Heterogeneous Integration

**Integration with Semiconductor Electronics**

Some interesting and potentially useful circuits have been demonstrated with thin-film BST technology, but the potential application space could be even greater if the films are successfully integrated with other technologies. Some significant progress has already been made in this direction. We have noted earlier that some of the initial motivation for developing thin-film BST was in connection with the silicon CMOS and DRAM industry [12]-[13], and processes were successfully developed there to integrate BST and STO films into Si integrated circuits as gate-oxides and DRAM capacitors [164]-[167].

One of the key integration challenges is the high processing temperature of BST and other related ceramic materials. Ordinarily the films are deposited and/or annealed at temperatures...
in the range of 600-700°C or higher, so the process must be sequenced in such a way that any temperature-sensitive steps are done after the ferroelectric deposition. Another challenge is that the tunable dielectrics are often grown in a highly oxidizing environment, which is necessary to minimize oxygen vacancies in the material. This oxidizing environment can adversely affect exposed semiconductor surfaces, so some kind of encapsulation layer is required to protect the epitaxial semiconductor layers. A process for integrating BST films with GaAs electronics has been reported [29] which addresses both of these concerns by first encapsulating the GaAs surface with silicon nitride, depositing and annealing the BST film, and finally removing the SiN encapsulant to fabricate the GaAs transistors and BST capacitors. Several circuits have been reported using this or a similar process [27,28], and it is believed that a substantial number of chip-sets have been sold that exploit this technology.

A similar process was recently developed for the integration of BST films with gallium nitride (GaN) electronics [168]. In this case a sacrificial SiO₂ layer was used to protect the the GaN epitaxial layers during BST growth. Figure 4-49 shows the measured I-V curves of device fabricated with and without this encapsulation layer, showing the importance of protecting the active layers during BST growth. Some progress has also been made in using BST as a gate oxide in GaN electronics to reduce gate leakage [169]. Figure 4-49 shows an example of an RF circuit that combines a GaN transistor and BST capacitor. In this particular circuit, the BST capacitor is only used as an AC bypass for bias decoupling, but results in a substantial size reduction (a SiN capacitor of similar capacity would occupy an area as large as the entire circuit of Figure 4-49).
**Integration with MEMS**

A final simple example of heterogeneous integration is combining high-\(\kappa\) materials like BST with RF Micro-ElectroMechanical Systems (MEMS). This is potentially interesting for two reasons. First, since BST is inherently an analog (continuously variable) varactor technology, there are certain circuit functions that can not be implemented easily using BST varactors, and hence there are a limited number of circuits that can be created using BST varactors alone. A low-loss and high-isolation switch is one example, something for which MEMS is particularly well-suited. Thus the combination of a MEMS switch and a BST varactor opens up a large number of new potential applications. Secondly, the MEMS devices themselves might benefit from the availability of a high-\(\kappa\) material like BST. For example, in the simple electrostatic switch shown in Figure 4-50, the dielectric layer plays an important role in determining the on-off capacitance ratio according to

\[
\frac{C_{\text{on}}}{C_{\text{off}}} = 1 + \varepsilon_r \frac{g}{d}
\]  

(4.46)

where \(d\) is the dielectric thickness. We can see that a high permittivity film can dramatically increase the on-off ratio and hence the performance of the device.

![Figure 4-50 – Simple MEMS electrostatic switch.](image)

This performance enhancement has been demonstrated in [171] and also recently in [172]. Figure 4-51 shows the improvement in switch isolation using a BST film in contrast to a more conventional SiN-based structure. There are some other technological issues such as dielectric breakdown strength that make the use of BST in this application somewhat challenging, but the example suggests some of the potential merits of marrying BST and MEMS technologies.

![Figure 4-51 – Simple MEMS electrostatic switch and improvement in down-state isolation using BST films [171].](image)
4.7 Conclusion and Acknowledgements

Tunable dielectric materials appear to have a strong potential for future wireless systems. The technology is cost-effective, compact, and with suitable device design can have excellent power-handling and linearity at low control voltages.

Presently the materials are reasonably well understood from a modeling and technology standpoint, and a variety of circuit demonstrations have been made. Phase-shifters, tunable matching networks, filters, phased-arrays, VCOs [173], switches [174], and a number of other proof-of-principle circuits have been reported, some of which have been reviewed here.

Much of the work presented in this chapter was based in whole or in part on work carried out at the University of California at Santa Barbara (UCSB) over the course of several years, and with the help of many students and collaborators. The author wishes to acknowledge several former students that were responsible for the development of this technology: Dr. Amit Nagra and Dr. Baki Acikel for their work in process development and distributed phase-shifter circuits; Dr. Troy Taylor and Dr. Peter Hansen for material growth and characterization; Dr. Yu Liu for MEMs integration; Justin Serraicco for work in compact phase-shifters; Dr. Hongtao Xu for RF device optimization and integration with GaN; and Dr. Nadia Pervez for work in the understanding and modeling of losses in BST capacitors. The author also acknowledges the contributions of Dr. James Speck and Dr. Susanne Stemmer in the Materials Dept. at UCSB, and many fruitful discussions with Dr. Stephen Streiffer at Argonne National Laboratories. Lastly, the authors would like to thank Dr. Chris Elsass, Dr. Vicki Chen, and Roger Forse at Agile Materials and Technologies Inc. in Goleta, CA for providing some of the materials and data used here.

Financial support for this work came from a variety of sources including: the Army Research Office through the Multifunctional Adaptive Radio Radar and Sensors program (MARRS MURI) DAAD19-01-1-0496; DARPA through the FAME (Frequency Agile Materials for Electronics) program, contract DABT63-98-1-0006; DMEA through the Center for Nanoscience Innovation for Defense (CNID), contract DMEA90-02-2-0215; and the Office of Naval Research through the Center for Advanced Nitride Electronics (CANE MURI), contract N00014-01-1-0764.

References


Bob York


Tunable Dielectrics for RF Circuits


[81] J. D. Baniecki, R. B. Laibowitz, T. M. Shaw, P. R. Duncombe, and D. A. Neumayer, D. E. Kotecki, H. Shen, Q. Y. Ma, “Dielectric relaxation of Ba0.7Sr0.3TiO3 thin films from 1 mHz to 20 GHz”, *Applied Physics Letters*, vol. 72, no. 4, 26 January 1998


Bob York


Tunable Dielectrics for RF Circuits


