Announcements & Agenda

- Project preliminary report due 11/21
- Multi-stage limiting amplifiers
- Bandwidth extension techniques
- Offset compensation
Limiting Amplifiers

- Limiting amplifier amplifies the TIA output to a reliable level to achieve a given BER with a certain decision element (comparator)

- Typically designed with a bandwidth of 1-1.2X data rate

- Want group delay variation $<\pm 10\%$ over bandwidth of interest to limit DDJ
How to Achieve an Amplifier GBW > $f_T$?

Assume for a 10Gb/s system that we need to build an amplifier with $A_v = 30dB$ and $f_{3dB} = 10GHz$.

$$GBW_{tot} = (31.6)(10GHz) = 316GHz$$

However, the peak $f_T$ of our technology is only 200GHz, and generally we can only achieve a single-stage amplifier GBW of

$$\text{Max Single-Stage } GBW_s \approx \frac{f_T}{3} = \frac{200GHz}{3} = 66.7GHz$$

with $A_v = 30dB \Rightarrow f_{3dB} = \frac{66.7GHz}{31.6} = 2.11GHz$, well below our 10GHz spec.

Instead of using a single-stage, let's break the amplifier into multiple stages with lower gain, but higher bandwidth. An optimal choice, from a maximum GBW perspective, is

$n = 7$ stages, with $A_{vs} = \sqrt[3]{31.6} = 1.64$ and $f_{3dB} = 31GHz$, or $GBW_s = 50.8GHz$

After multi-stage bandwidth compression, this will yield a total GBW $\approx 316GHz$ with our target gain of 31.6 and with a single-stage $GBW_s = 50.8GHz$ that our technology can support.
Multi-Stage Amplifier GBW

If every stage is a single-pole amplifier

\[
\frac{A_{vs}}{1 + \frac{s}{\omega_{3dBs}}} = \frac{A_{vs}^n}{1 + \frac{s}{\omega_{3dBs}}}^{\frac{n}{n}}
\]

The total multi-stage amplifier transfer function will be

\[
\frac{v_{out}}{v_{in}} = \left( \frac{A_{vs}}{1 + \frac{s}{\omega_{3dBs}}} \right)^n = A_{vs}^n \left( 1 + \frac{s}{\omega_{3dBs}} \right)^n
\]

The gain has increased significantly, but the bandwidth does compress relative to a single stage.
The total amplifier 3-dB bandwidth, $\omega_{3\text{dBlot}}$, is where

\[
\frac{v_{\text{out}}}{v_{\text{in}}} = \left| \frac{A}{1 + \frac{j\omega_{3\text{dBlot}}}{\omega_{3\text{dbs}}}} \right|^n = \frac{A^n}{\sqrt{2}}
\]

\[
\left( \frac{A_{vs}}{\sqrt{1 + \left( \frac{\omega_{3\text{dBlot}}}{\omega_{3\text{dbs}}} \right)^2}} \right)^n = \frac{A_{vs}^n}{\sqrt{2}}
\]

\[
\left( 1 + \left( \frac{\omega_{3\text{dBlot}}}{\omega_{3\text{dbs}}} \right)^2 \right)^n = 2
\]

\[
\omega_{3\text{dBlot}} = \omega_{3\text{dbs}} \sqrt{\frac{1}{2^n} - 1}
\]

The total multi-stage bandwidth does compress, although at a much slower rate than the increase in gain.

Thus, a significant increase in GBW can be achieved with a multi-stage amplifier approach.
Optimum Number of Gain Stages

Assuming that there is a maximum per-stage $GBW_s$ that the technology can support

$$GBW_s = A_{vs} \omega_{3dBs} \Rightarrow \omega_{3dBs} = \frac{GBW_s}{A_{vs}}$$  (Note, here GBW is in rad/s)

If we need to achieve a high bandwidth, we have to reduce the per-stage gain and increase the number of stages. However, the bandwidth will compress with cascaded stages. Thus, there must be an optimum number of stages for a maximum potential gain bandwidth.

Recall that the total bandwidth is

$$\omega_{3dBtot} = \omega_{3dBs} \sqrt{2^{\frac{1}{n}} - 1} = \frac{GBW_s}{A_{vs}} \sqrt{2^{\frac{1}{n}} - 1}$$

and we will achieve a total gain $G_{tot}$ with $n$ stages

$$A_{vs} = G_{tot}^{\frac{1}{n}} \Rightarrow \omega_{3dBtot} = \frac{GBW_s}{G_{tot}^{\frac{1}{n}}} \sqrt{2^{\frac{1}{n}} - 1}$$
Optimum Number of Gain Stages

For a given total gain, we would like to maximize the bandwidth. In order to do this, let's make the following approximation

\[
\omega_{3dB_{tot}} = \frac{GBW_s}{G_{tot}^{1/n}} \sqrt{2^{\frac{1}{n}} - 1} \approx \frac{GBW_s}{G_{tot}^{1/n}} \sqrt{\frac{1}{n} \ln 2}
\]

Also, instead of maximizing this expression, let's minimize its reciprocal w.r.t the number of stages

\[
\frac{1}{\omega_{3dB_{tot}}} = \left( \frac{\sqrt{n}}{GBW_s \sqrt{\ln 2}} \right) G_{tot}^{1/n}
\]

\[
\frac{d}{dn} \left( \frac{1}{\omega_{3dB_{tot}}} \right) = \frac{d}{dn} \left( \left( \frac{\sqrt{n}}{GBW_s \sqrt{\ln 2}} \right) G_{tot}^{1/n} \right) = 0
\]

Moreover, to make this easier, let's minimize the natural log of the denominator, as this should yield the same optimum.

\[
\frac{d}{dn} \left( \ln \left( \frac{1}{\omega_{3dB_{tot}}} \right) \right) = \frac{d}{dn} \left( \ln \left( \left( \frac{\sqrt{n}}{GBW_s \sqrt{\ln 2}} \right) G_{tot}^{1/n} \right) \right) = \frac{d}{dn} \left( \frac{1}{2} \ln(n) + \frac{1}{n} \ln(G_{tot}) - \ln(GBW_s \sqrt{\ln 2}) \right) = 0
\]
Optimum Number of Gain Stages

\[
\frac{d}{dn}\left(\ln\left(\frac{1}{\omega_{3dB_{tot}}}\right)\right) = \frac{d}{dn}\left(\ln\left(\frac{\sqrt{n}}{\sqrt[4]{GBW_s\ln 2}}G_{tot}^{1/n}\right)\right) = \frac{d}{dn}\left(\frac{1}{2}\ln(n) + \frac{1}{n}\ln(G_{tot}) - \ln(GBW_s\ln 2)\right) = 0
\]

\[
\frac{1}{2n} - \frac{1}{n^2}\ln(G_{tot}) = 0
\]

\[
\frac{1}{n}\ln(G_{tot}) = \frac{1}{2}
\]

Thus, the optimum number of stages is

\[
n_{opt} = 2\ln(G_{tot})
\]

and the optimum stage gain is

\[
A_{vs, opt}^{2\ln(G_{tot})} = G_{tot}
\]

\[
2\ln(G_{tot})\ln(A_{vs, opt}) = \ln(G_{tot})
\]

\[
A_{vs, opt} = \sqrt{e} = 1.65
\]
Optimum Number of Gain Stages

For example, a multi-stage amplifier with $G_{tot} = 100$ should have

$$n_{opt} = 2 \ln (G_{tot}) = 2 \ln (100) = 9.21$$

Assuming 9 stages results in

$$A_{vs} = \sqrt[9]{100} = 1.67$$

which is close to $\sqrt{e} = 1.65$

Relative to the per-stage bandwidth, the total amplifier bandwidth will compress to

$$\omega_{3dB_{tot}} = \omega_{3dBs} \sqrt{2^{\frac{1}{9}} - 1} = 0.283 \omega_{3dBs}$$

- Note, while this is the optimum number of stages from a maximum GBW perspective, the bandwidth doesn’t falloff too dramatically with lower $n$

- Thus, from a power and noise perspective, it may make sense to use a lower number of LA stages

- Typically high-gain LAs use between 3-7 stages
Bandwidth Extension Techniques

• In order to increase the bandwidth of our multi-stage amplifiers, we need to increase the bandwidth of the individual stages

• Passive bandwidth extension techniques
  • Shunt Peaking
  • Series Peaking
  • T-coil Peaking

• An excellent reference
Shunt Peaking

• Adding an inductor in series with the load resistor introduces a zero in the impedance transfer function

• This zero increases the impedance with frequency, compensating the decrease caused by the capacitor, and extending the bandwidth

\[ Z(s) = \frac{V_{out}}{I_{in}} = \left( \frac{1}{sC} \right) \| (R+sL) = \frac{R + sL}{1 + sRC + s^2LC} \]
Shunt Peaking

- While the inductor can increase the bandwidth significantly, frequency peaking can occur if the inductor is too big.
- For a flat frequency response, ~70% bandwidth increase can be achieved.
- A maximum 85% bandwidth increase is possible with 1.5dB of peaking.

![Graph showing frequency response with and without shunt peaking.](image)

**Table: Ratio of time constants and normalized bandwidths**

<table>
<thead>
<tr>
<th>Condition</th>
<th>Ratio of time constants</th>
<th>Normalized bandwidth</th>
<th>Normalized peak frequency response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum bandwidth</td>
<td>( m = R^2C/L )</td>
<td>( \sim 1.41 )</td>
<td>( \sim 1.85 )</td>
</tr>
<tr>
<td>(</td>
<td>Z</td>
<td>= R \ @ \ \omega = 1/RC)</td>
<td></td>
</tr>
<tr>
<td>Maximally flat frequency response</td>
<td>( \sim 2.41 )</td>
<td>( \sim 1.72 )</td>
<td>1</td>
</tr>
<tr>
<td>Best group delay</td>
<td>( \sim 3.1 )</td>
<td>( \sim 1.6 )</td>
<td>1</td>
</tr>
<tr>
<td>No shunt peaking</td>
<td>( \infty )</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Bridged-Shunt Peaking

Adding a bridge capacitor in parallel with the inductor allows for compensation of the frequency peaking with the possible maximum shunt peaking bandwidth increase.

\[
Z_N(s) = \frac{1 + \left( \frac{1}{m} \right) \frac{s}{\omega_0} + \left( \frac{k_B}{m} \right) \frac{s^2}{\omega_0^2}}{1 + \frac{s}{\omega_0} + \left( \frac{k_B + 1}{m} \right) \frac{s^2}{\omega_0^2} + \left( \frac{k_B}{m} \right) \frac{s^3}{\omega_0^3}}
\]

\[
k_B = C_B / C, \omega_0 = 1/RC, \text{ and } m = R^2C/L
\]
Series Peaking

- Introducing a series peaking inductor is useful to “split” the load capacitance between the amplifier drain capacitance and the next stage gate capacitance.
- Without L, the transistor has to charge the total capacitance at the same time.
- With L, initially only $C_1$ is charged, reducing the risetime at the drain and increasing bandwidth.

\[ Z_N(s) = \frac{1}{1 + \frac{s}{\omega_0} + \left( \frac{1 - k_C}{m} \right) \frac{s^2}{\omega_0^2} + \left( \frac{k_C(1 - k_C)}{m} \right) \frac{s^3}{\omega_0^3}} \]

\[ k_C = \frac{C_1}{C} \quad m = \frac{R^2C}{L} \]
Series Peaking

- As the capacitance is more distributed with a higher $k_c$ value, a higher BWER is achieved.
- Up to 1.5x bandwidth increase is achieved with no peaking.
- Higher BWER is possible with some frequency peaking.

<table>
<thead>
<tr>
<th>$k_c = \frac{C_1}{C}$</th>
<th>Ripple (dB)</th>
<th>$m = R^2C/L$</th>
<th>BWER</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>2</td>
<td>1.41</td>
</tr>
<tr>
<td>0.1</td>
<td>0</td>
<td>1.8</td>
<td>1.58</td>
</tr>
<tr>
<td>0.2</td>
<td>0</td>
<td>1.8</td>
<td>1.87</td>
</tr>
<tr>
<td>0.3</td>
<td>0</td>
<td>2.4</td>
<td>2.52</td>
</tr>
<tr>
<td>0.4</td>
<td>1</td>
<td>1.9</td>
<td>2.75</td>
</tr>
<tr>
<td>0.5</td>
<td>3.3</td>
<td>1.5</td>
<td>2.65</td>
</tr>
</tbody>
</table>
Bridged-Shunt-Series Peaking

- Combining both shunt and series peaking can yield even higher bandwidth extension

\[ Z_N(s) = \frac{1 + \left( \frac{1}{m_1} \right) \frac{s}{\omega_0} + \left( \frac{k_B}{m_1} \right) \frac{s^2}{\omega_0^2}}{1 + \frac{s}{\omega_0} + \left( \frac{1+k_B}{m_1} + \frac{1-k_C}{m_2} \right) \frac{s^2}{\omega_0^2} + \left( \frac{k_B}{m_1} + \frac{k_C(1-k_C)}{m_2} \right) \frac{s^3}{\omega_0^3} + \left( \frac{(k_C+k_B)(1-k_C)}{m_1 m_2} \right) \frac{s^4}{\omega_0^4} + \left( \frac{k_B k_C (1-k_C)}{m_1 m_2} \right) \frac{s^5}{\omega_0^5}} \]

\[ m_1 = \frac{R^2 C}{L_1} \quad m_2 = \frac{R^2 C}{L_2} \]
Bridged-Shunt-Series Peaking

Proper choice of component values can yield close to 4x increase in bandwidth with no peaking.

However, this requires tight control of these components, which can be difficult with PVT variations.
T-Coil Peaking

- If the input transistor drain capacitance ($C_1$) is relatively small, then the bandwidth extension through shunt-series peaking is limited.

- T-coil peaking, which utilizes the magnetic coupling of a transformer, provides better bandwidth extension in this case:
  - $L_2$ performs capacitive splitting, such that the initial current charges only $C_1$
  - As current begins to flow through $L_2$, magnetically coupled current also flows through $L_1$, providing increased current to charge $C_2$ which improves bandwidth and transition times.
T-Coil Peaking

\[
Z_N(s) = \frac{1 + \left( \frac{1}{m_1} + \frac{k_m}{\sqrt{m_1 m_2}} \right) \frac{s}{\omega_0}}{1 + \frac{s}{\omega_0} + \left( \frac{1}{m_1} + \frac{k_C m_2}{\sqrt{m_1 m_2}} \right) \frac{s^2}{\omega_0^2} + \left( \frac{k_C (1 - k_C)}{m_2} \right) \frac{s^3}{\omega_0^3} + \left( \frac{k_C (1 - k_C) (1 - k_m^2)}{m_1 m_2} \right) \frac{s^4}{\omega_0^4}}
\]

\[k_m = \frac{M}{\sqrt{L_1 L_2}}.\]
T-Coil Peaking

- A bandwidth extension of 4x is possible without any frequency peaking
- If peaking is acceptable, then a BWER near 5 can be achieved, depending on the size of $C_1$

<table>
<thead>
<tr>
<th>$k_C = C_2 / C$</th>
<th>Ripple (dB)</th>
<th>$m_1 = R_1 C_1 L_1$</th>
<th>$m_2 = R_2 C_1 L_2$</th>
<th>$k_m = M /\sqrt{I_1 I_2}$</th>
<th>BWER</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>0</td>
<td>4</td>
<td>1.6</td>
<td>-0.7</td>
<td>4.63</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>3.5</td>
<td>1.2</td>
<td>-0.6</td>
<td>4.92</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>3.5</td>
<td>1.6</td>
<td>-0.6</td>
<td>5.59</td>
</tr>
<tr>
<td>0.2</td>
<td>0</td>
<td>5.5</td>
<td>2.4</td>
<td>-0.6</td>
<td>4.14</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>-0.6</td>
<td>4.51</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>2.4</td>
<td>-0.5</td>
<td>4.86</td>
</tr>
<tr>
<td>0.3</td>
<td>0</td>
<td>4</td>
<td>2.8</td>
<td>-0.5</td>
<td>3.93</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>3.5</td>
<td>2</td>
<td>-0.4</td>
<td>3.98</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
<td>2.8</td>
<td>-0.4</td>
<td>4.54</td>
</tr>
</tbody>
</table>
Active Bandwidth Extension Techniques

• While passive techniques offer excellent bandwidth extension at near zero power cost, there are some disadvantages
  • Generally large area
  • Process support/characterization of inductors/transformers

• Active circuit techniques can also be employed to extend amplifier bandwidth

• Some active bandwidth extension techniques
  • Negative Miller Capacitance
  • Active Negative Feedback

• There are numerous other techniques, but that is all we have time for this semester 😊
Negative Miller Capacitance

- In modern technologies, Cgd is a significant (50% to near 100%) fraction of Cgs
- Amplifier effective input capacitance can increase significantly due to the Miller multiplication of Cgd
- Without additional Cn:

\[ C_{in} = C_{gs1} + C_{gd1}(1 - A_{gd}) \]

As \( A_{gd} \) is negative, and often is the differential gain of the amplifier, this can result in significant increase in the effective input capacitance.
Negative Miller Capacitance

- In order to mitigate this Cgd multiplication, additional cross-coupled capacitors can be added from the amplifier inputs to the outputs.

- Effectively, the charge on this additional capacitor charges a (large) portion of the Cgd capacitor.

\[
C_{in} = C_{gs1} + C_{gd1} (1 - A_{gd}) + C_n (1 - (-A_{gd}))
\]

If \(C_n\) is set equal to \(C_{gd1}\)

\[
C_{in} = C_{gs1} + 2C_{gd1}
\]

Thus, as long as the amplifier gain is > 1, a reduction in the effective input capacitance is achieved.
Active Negative Feedback

- Instead of using simple first-order amplifier cells, a second-order cell with active negative feedback can provide bandwidth enhancement.

\[ \frac{V_{\text{out}}}{V_{\text{in}}} = \frac{A_{vo}\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \]

\[ A_{vo} = \frac{G_{m1}G_{m2}R_{L1}R_{L2}}{1 + G_{m2}G_{mf}R_{L1}R_{L2}} \]

\[ \zeta = \frac{1}{2} \frac{R_{L1}C_1 + R_{L2}C_2}{\sqrt{R_{L1}R_{L2}C_1C_2(1 + G_{mf}G_{m2}R_{L1}R_{L2})}} \]

\[ \omega_n^2 = \frac{1 + G_{mf}G_{m2}R_{L1}R_{L2}}{R_{L1}R_{L2}C_1C_2}. \]
Active Negative Feedback

- This second-order amplifier cell can be optimized for different objectives, but $G_m f$ can be set to yield a Butterworth response with a maximally-flat frequency response

\[
\zeta = \sqrt{2}/2
\]

\[
\omega_{-3dB} = 2\pi f_{-3dB} = \omega_n/(2\pi)
\]

\[
A_{vo}\omega_{-3dB}^2 = \frac{G_{m1} G_{m2}}{C_1 C_2}
\]

\[
A_{vo}\omega_{-3dB} = \frac{G_{m1} G_{m2}}{C_1 C_2} \frac{1}{\omega_{-3dB}}.
\]
Active Negative Feedback

\[ A_{vo \omega -3dB} = \frac{G_{m1} G_{m2}}{C_1 C_2} \frac{1}{\omega -3dB}. \]

The ratio \( \frac{G_m}{C} \) is proportional to the technology \( \omega_T \)

\[ \frac{G_m}{C} = \alpha \omega_T \]

Assuming \( \frac{G_{m1}}{C_1} \approx \frac{G_{m2}}{C_2} \approx \alpha \omega_T \)

\[ A_{vo \omega -3dB} = \frac{(\alpha \omega_T)^2}{\omega -3dB} = \alpha^2 \omega_T \frac{\omega_T}{\omega -3dB} \]

- The second-order cell gain-bandwidth can potentially achieve a value greater than the technology \( f_T \)
Limiting Amplifier Example 1

[Galal JSSC 2003]

- Resistive Load Only
- Active Negative Feedback
- Shunt Inductive Peaking
- Negative Miller Capacitance
Limiting Amplifier Example 2

- T-coils in LA stages allow for a combination of series and shunt peaking and close to 3x bandwidth extension

[Proesel ISSCC 2012]
Offset Compensation

- The receiver sensitivity is degraded if the limiting amplifier has an input-referred offset.
- This is often quantified in terms of a Power Penalty, PP:

\[ PP = \frac{v_{l_{pp}} + 2V_{OS}}{v_{l_{pp}}} = 1 + \frac{2V_{OS}}{v_{l_{pp}}} \]

- It is important to minimize the offset of these multi-stage limiting amplifiers!
Offset Compensation

- The DC offset, $V_{os}$, of the limiting amplifier is compensated by a low-frequency negative feedback loop.

Ideally, this reduces the offset to

$$\frac{V_{os}}{AA_l}$$

However, if the error amplifier has an offset, $V_{os1}$, the offset becomes

$$\sqrt{\left(\frac{V_{os}}{AA_l}\right)^2 + \left(\frac{V_{os1}}{A}\right)^2}$$
Offset Compensation

The low-pass filtering in the feedback loop causes a low-frequency cutoff

\[
f_{LF} = \frac{1}{2\pi} \frac{AA_i/2+1}{R_iC_1}
\]

Thus, the feedback loop bandwidth should be made much lower than the lowest frequency content of the input data.

This may lead to large-area passive in the offset correction feedback.

Some designs leverage Miller capacitive multiplication with the error amplifier to reduce this filter area.
Next Time

- High-Speed Transmitters