Blind Digital Tuning for an Analog World
- Radio Self-Interference Cancellation

Yingbo Hua
University of California at Riverside

July 2015
Global mobile data traffic grew 69 percent in 2014... (Cisco).

Demand for radio spectrum continues to increase.

Spectrum sharing between communications and passive radar becomes necessary.

The best radio transceivers in the future must be able to transmit and receive using the same frequency at the same time (*full-duplex radio*).

*Radio self-interference cancellation* will enable full-duplex radio and facilitate spectrum sharing between communications and radar.
In addition to antenna/circuit isolation, radio self-interference cancellation must be done first at radio frontend.
The interference channel’s impulse response in baseband:

\[ h_{\text{Int}}(t) = \sum_{i=1}^{l} a_i e^{-2\pi f_c \tau_i} \text{sinc}(W(t - \tau_i)) \]

How do we choose an analog cancellation channel to match the above by using only passive components with minimum noise?
An Analog Cancellation Channel

- The RF waveform only passes through passive components.
- The phase shifters and time delays are constant with possible constant errors.
- The only variable components are step-attenuators which are digitally controlled.
Attenuation in polar form, i.e., $-20 \log |G| e^{-\text{arg}(G)}$

**Figure:** 1 attenuator per c-tap

**Figure:** 2 attenuators per c-tap
Attenuation in polar form, i.e., $-20 \log |G| e^{-j\arg(G)}$

**Figure:** 3 attenuators per c-tap

**Figure:** 4 attenuators per c-tap
An Alternative Form

- Using 90-degree power splitters to implement desired phases.
Comparison of Ideal Channel Models:

\[ h_{\text{Int}}(t) = \sum_{i=1}^{l} a_i e^{-2\pi f_c \tau_i} \text{sinc}(W(t - \tau_i)) \]

\[ h_{\text{Can}}(t) = e^{-j2\pi f_c T_0} \sum_{n=0}^{N-1} G_n \text{sinc}(W(t - nT)) \]

- \( |G_n| = |g_{n,1} + jg_{n,2} - g_{n,3} - jg_{n,4}| < 1 \)
- We want \( h_{\text{Res}}(t) = h_{\text{Int}}(t) - h_{\text{Can}}(t) \) to be as small as possible by adjusting all \( g_{n,m} \).
- If \( \max_{i,j} |\tau_i - \tau_j| \ll \frac{1}{W} \), \( N = 1 \) would be sufficient.
- Otherwise, \( N > 1 \) is needed.
Minimized Residual Self-Interference Channel

Figure: The distribution of the residual interference over $f$, i.e., $E_2(f_m)$. $T = \frac{1}{10W}$. $\alpha_p = 2$.

Figure: The CDF of the residual interference, i.e., the CDF of $E_1^{(r)}$. $T = \frac{1}{10W}$. $\alpha_p = 2$.

\[ E_2(f) = E\{ \left| \frac{H_{Res}(f)}{H_{Int}(f)} \right|^2 \} \] and $E_1^{(r)}$ is the $r$th run of

\[ \int_{-W/2}^{W/2} \left| \frac{H_{Res}(f)}{H_{Int}(f)} \right|^2 df \]
Practical Constraints

- $h_{Int}(t)$ is unknown except for some poor estimates.
- $h_{Can}(t)$ is also an unknown function of the gains/attenuations of the step attenuators due to analog interface and hardware imperfection.
- There is no precise access to the input and/or output of either channel.
- A step attenuator has a limited dynamic range and a fixed step size.
- To tune a practical cancellation channel, we must tolerate many unknown factors - *blind tuning*.
- Step attenuators allow *blind digital tuning*.
Each c-tap has four independent step-attenuators. The complex attenuation of a c-tap densely covers four quadrants of a disk.

Each step-attenuator has a step size 1dB and 1-32dB dynamic range. There are $32^4N$ choices for a cancellation channel with $N$ c-taps - too slow to do brute force blind digital tuning.
Methods to Speed Up Blind Digital Tuning:

- Using Additional Down-Converters - Method 1
- Using No Additional Down-Converter - Method 2
- Using Quadratic Model - Method 3
Using Additional Down-Converters - Method 1

- Assuming that we can introduce a down-converter to tap the input RF waveform $x_i(k)$ of each step-attenuator, then ideally we can write the residual interference $x_{res}(k)$ measurable in baseband as

$$x_{res}(k) = x_0(k) - T\{[x_1(k), \cdots, x_{NA}(k)][g_1, \cdots, g_{NA}]^T\}$$

- Here $T$ is some unknown linear operator caused by unknown RF couplings.

- If $T$ is known and $x_i(k)$ can be measured accurately enough, the classic LS, RLS and LMS algorithms can be applied.

- But in practice, $T$ is unknown and $x_i(k)$ is contaminated by noise. The down-converters are also costly.
Without any additional down-converter, we can write the residual interference measurable in baseband as

\[
x_{\text{res}}(k) = x_0(k) - X(k)g
\]

where \(x_0(k)\) and \(X(k)\) are unknown.

The affine/linear relationship holds strongly (in case of IQ imbalances) if a modification into real-valued representation is used.

With unknown \(x_0(k)\) and \(X(k)\), how to find the optimal \(g\) to minimize \(x_{\text{res}}(k)\)?
Define a sequence of training vectors of \( g \): i.e., \( g_1, \cdots, g_{N_T} \).

For each training vector \( g_i \), Tx transmits the same data multiple times, and Rx measures the averaged residual \( \bar{x}_{\text{res},i}(k) \), which can be written as

\[
\bar{x}_{\text{res},i}(k) = \bar{x}_0(k) - \bar{X}(k)g_i + w_i(k), \; i = 1, \cdots, N_T
\]

where \( \bar{x}_0(k) \) and \( \bar{X}(k) \) can be treated as independent of \( i \).

Provided that

\[
\begin{bmatrix}
1 & \cdots & 1 \\
g_1 & \cdots & g_{N_T}
\end{bmatrix}
\]

has a full row rank, there is a unique LS solution for \( \bar{x}_0(k) \) and \( \bar{X}(k) \).

Then, the optimal \( g \) follows.
The previous affine model is sensitive to phase noise.

One way to remove the phase noise is to choose the power of $x_{res}(k)$ as the measurement, which leads to

$$p_{res} = g^T A g + g^T b + c$$

where $A$, $b$ and $c$ are unknown parameters.

How to find the optimal $g$ to minimize $p_{res}$?
If the estimates $\hat{A}$ and $\hat{b}$ are given, the optimal $g$ is

$$\hat{g} = -\frac{1}{2}\hat{A} + \hat{b}$$

To find $A$ and $b$, we can first define a sequence of training vectors of $g$: i.e., $g_1, \cdots, g_{N_T}$ (not the same as before).

For each $g_i$, Tx transmits a stream of data, and Rx measures the residual power $p_{\text{res},i}$. We can then write

$$\begin{bmatrix} p_{\text{res},1} \\ \vdots \\ p_{\text{res},N_T} \end{bmatrix} = \begin{bmatrix} 1 & g_1^T & g_1^T \otimes g_1^T \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & g_{N_T}^T & g_{N_T}^T \otimes g_{N_T}^T \end{bmatrix} \begin{bmatrix} c \\ b \\ \text{vec}\{A\} \end{bmatrix} = G\theta$$

If the training matrix $G$ had a full column rank, there would be a unique LS solution for $\theta$. 
The Quadratic Term $g^T A g$

With $g$ being real, the unknown matrix $A$ in $g^T A g$ can always be chosen to be real symmetric. Hence, we can write

$$g^T A g = (g^T \otimes g^T) \text{vec}\{A\} = (g^T \otimes g^T) S^T D S \text{vec}\{A\}$$

where, assuming $g$ being $4 \times 1$ and $J_{M \times N}$ being the last $M$ rows of $I_{N \times N}$,

$$S = \text{diag}[J_{4 \times 4}, J_{3 \times 4}, J_{2 \times 4}, J_{1 \times 4}]$$

$$D = \text{diag}[1, 2, 2, 2, 1, 2, 2, 1, 2, 1]$$
Hence, we can rewrite $\mathbf{G}\theta$ as

$$\begin{bmatrix}
  p_{\text{res},1} \\
  \vdots \\
  p_{\text{res},N_T}
\end{bmatrix}
= \begin{bmatrix}
  1 & \mathbf{g}_1^T & (\mathbf{g}_1^T \otimes \mathbf{g}_1^T)\mathbf{S}^T \\
  \vdots & \vdots & \vdots \\
  1 & \mathbf{g}_{N_T}^T & (\mathbf{g}_{N_T}^T \otimes \mathbf{g}_{N_T}^T)\mathbf{S}^T
\end{bmatrix}
\begin{bmatrix}
  \mathbf{c} \\
  \mathbf{b}
\end{bmatrix}
\, \text{DSvec}\{\mathbf{A}\}$$

or equivalently

$$\mathbf{p} = \mathbf{G}_T\theta_T$$

Here, $\mathbf{G}_T$ and $\theta_T$ of $N_T \times (\frac{1}{2}N_A(N_A + 1) + N_A + 1)$ and $(\frac{1}{2}N_A(N_A + 1) + N_A + 1) \times 1$ are smaller than $\mathbf{G}$ and $\theta$.

For a unique LS solution of $\mathbf{A}$ and $\mathbf{b}$, we only need $\mathbf{G}_T$ to be of full column rank.
Example of $G_T$ of Full Column Rank

- Assume $g_1 = 0$ and the following:
  - For $i = 2, \cdots, N_A + 1$, $g_i = \alpha e_{N_A, i-1}$ with $0 < \alpha \leq 1$ and $e_{N_A, i}$ being the $i$th column of $I_{N_A \times N_A}$
  - For $i = N_A + 1, \cdots, 2N_A + 1$, $g_i = \beta e_{N_A, i-N_A-1}$ with $0 < \beta < 1$ and $\beta < \alpha$
  - For $i = 2N_A + 2, \cdots, \frac{1}{2}(N_A + 1)N_A + N_A + 1$, $g_i = \alpha e_{N_A, l} + \alpha e_{N_A, k}$ with $1 \leq l < k \leq N_A$

- Then, we can show that

$$\det\{G_T\} = \alpha^{N_A^2} \beta^{N_A} (\alpha - \beta)^{N_A} \neq 0$$

- The above matrix $G_T$ is very sparse, and a closed-form expression of $G_T^{-1}$ can also be found.
Example of $G_T$ with $N_A = 3$

$$\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & \alpha & 0 & 0 & \alpha^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & \beta & 0 & 0 & \beta^2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & \alpha & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 & 0 & 0 \\
1 & 0 & \beta & 0 & 0 & 0 & 0 & \beta^2 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & \alpha^2 & 0 & 0 \\
1 & 0 & 0 & \beta & 0 & 0 & 0 & 0 & 0 & 0 & \beta^2 & 0 \\
1 & \alpha & \alpha & 0 & \alpha^2 & \alpha^2 & 0 & \alpha^2 & 0 & 0 & 0 & 0 \\
1 & \alpha & 0 & \alpha & \alpha^2 & 0 & \alpha^2 & 0 & 0 & 0 & \alpha^2 & 0 \\
1 & 0 & \alpha & \alpha & 0 & 0 & 0 & \alpha^2 & \alpha^2 & \alpha^2 & \alpha^2 \\
\end{bmatrix}$$
Example of Solution to $p = G_T\theta_T$

$$Hp = P\theta_T$$

Figure: The matrix $H$ with
\[ \Delta = \alpha \beta^2 - \alpha^2 \beta, \ \Delta_1 = \alpha^2 - \beta^2 \]
and \[ \Delta_2 = \beta - \alpha. \]

Figure: The permutation matrix $P$. 
Basic Facts of Radio Hardware

- Phase Noise
- IQ Imbalance
- Real Linear Model vs Widely Linear (Complex) Model
Phase Noise

- **Tx phase noise**: for a baseband signal $x_{BB}(t) = x_r(t) + jx_i(t)$, its RF signal generated by a practical transmitter is

$$x_{RF}(t) = x_r(t) \cos(2\pi f_c t + \phi_T(t)) - x_i(t) \sin(2\pi f_c t + \phi_T(t))$$

where $\phi_T(t)$ is the Tx phase noise.

- **Rx phase noise**: for the above RF signal, its baseband signal generated by a practical receiver is

$$\hat{x}_{BB}(t) = x_{BB}(t)e^{j\phi_T(t)+j\phi_R(t)}$$

where $\phi_R(t)$ is the Rx phase noise.

- Phase noise is non-additive noise. If approximated via Taylor’s series, the noise term is correlated with the signal.
IQ Imbalances

▶ Tx IQ imbalances: for a baseband signal
\[ x_{BB}(t) = x_r(t) + jx_i(t), \]
the RF signal generated by a practical transmitter is
\[ x_{RF}(t) = (1+\delta_T)x_r(t)\cos(2\pi f_c t + \theta_T) - (1-\delta_T)x_i(t)\sin(2\pi f_c t - \theta_T) \]
where \( \delta_T \) and \( \theta_T \) are the Tx amplitude and phase imbalances.

▶ Rx IQ imbalances: for the above RF signal, the baseband signal generated by a practical receiver is
\[ \hat{x}_{BB}(t) = \hat{x}_r(t) + j\hat{x}_i(t) \]
and
\[
\begin{bmatrix}
\hat{x}_r(t) \\
\hat{x}_i(t)
\end{bmatrix} =
\begin{bmatrix}
(1 + \delta_R)\cos \theta_R & (1 + \delta_R)\sin \theta_R \\
(1 - \delta_R)\sin \theta_R & (1 - \delta_R)\cos \theta_R
\end{bmatrix}
\begin{bmatrix}
(1 + \delta_T)\cos \theta_T & (1 - \delta_T)\sin \theta_T \\
(1 + \delta_T)\sin \theta_T & (1 - \delta_T)\cos \theta_T
\end{bmatrix}
\begin{bmatrix}
x_r(t) \\
x_i(t)
\end{bmatrix}
\]

Rx IQ imbalance
\[ \begin{bmatrix}
\hat{x}_r(t) \\
\hat{x}_i(t)
\end{bmatrix} =
\begin{bmatrix}
(1 + \delta_R)\cos \theta_R & (1 + \delta_R)\sin \theta_R \\
(1 - \delta_R)\sin \theta_R & (1 - \delta_R)\cos \theta_R
\end{bmatrix}
\begin{bmatrix}
x_r(t) \\
x_i(t)
\end{bmatrix}
\]
TX IQ imbalance

▶ The real-valued linear model holds but the complex-valued linear model does not.
With IQ imbalances, a MIMO radio channel can be described by

\[ y_{re}(k) = \sum_{l} H_{re}(l)x_{re}(k - l) + w_{re}(k) \]

or equivalently by

\[ y_{co}(k) = \sum_{l} A_{co}(l)x_{co}(k - l) + \sum_{l} B_{co}(l)x_{co}^*(k - l) + w_{co}(k) \]

The real linear model is clearly more straightforward to use, e.g., transmit beamforming designing.

Why is the widely linear model “more popular”? 
Assume that we want to design a MIMO zero-forcing beamformer for $n_t$ transmitters, which produces nulls at $n_r$ ($<n_t$) receivers.

The Tx waveform should be $x_{re}(k) = P_{re}(k) * s_{re}(k)$ and

$$H_{re}(k) * P_{re}(k) = \begin{bmatrix} H_{re,a}(k) & H_{re,b}(k) \end{bmatrix} \begin{bmatrix} P_{re,a}(k) \\ 2n_r \times 2n_r \\ H_{re,a}(k) & H_{re,b}(k) \end{bmatrix} = 0$$

A solution is $P_{re,a} = \text{Adj}\{H_{re,a}(k)\} * H_{re,b}(k)$ and $P_{re,b}(k) = -\text{Det}\{H_{re,a}(k)\} I_{2(n_t-n_r) \times 2(n_t-n_r)}$

The beamformer for widely linear model can be found but needs extra work.
Example: Channel Estimation

- **Real linear model:**

\[
\begin{bmatrix}
  y_{re}(0), \ldots, y_{re}(N-1)
\end{bmatrix}_{R^{2n_r \times N}} = \begin{bmatrix}
  H_{re}(0), \ldots, H_{re}(L)
\end{bmatrix}_{R^{2n_r \times N}}
\]

\[
\begin{bmatrix}
  x_{re}(0) & \ldots & x_{re}(N-1) \\
  x_{re}(-L) & \ldots & x_{re}(N-1-L)
\end{bmatrix}_{x_{re} \in R^{2n_t(L+1) \times N}}
\]

It requires \((X_{re}X_{re}^T)^{-1} \in R^{2n_t(L+1) \times 2n_t(L+1)}\)

- **Widely linear model**

\[
\begin{bmatrix}
  y_{co}(0), \ldots, y_{co}(N-1)
\end{bmatrix}_{C^{n_r \times N}} = \begin{bmatrix}
  A_{co}(0), \ldots, A_{co}(L)
\end{bmatrix}_{A_{co} \in C^{n_r \times N}}
\]

\[
\begin{bmatrix}
  x_{co}(0) & \ldots & x_{co}(N-1) \\
  x_{co}(-L) & \ldots & x_{co}(N-1-L)
\end{bmatrix}_{x_{co} \in C^{n_t(L+1) \times N}}
\]

\[
+ \begin{bmatrix}
  B_{co}(0), \ldots, B_{co}(L)
\end{bmatrix}_{B_{co} \in C^{n_r \times N}}
\]

\[
\begin{bmatrix}
  x_{co}^*(0) & \ldots & x_{co}^*(N-1) \\
  x_{co}^*(-L) & \ldots & x_{co}^*(N-1-L)
\end{bmatrix}_{x_{co}^* \in C^{n_t(L+1) \times N}}
\]

It requires \(\left(\begin{bmatrix}
  X_{co}^* & X_{co}^* \end{bmatrix} \begin{bmatrix}
  X_{co} & X_{co}^* \end{bmatrix}^H\right)^{-1} \in C^{2n_t(L+1) \times 2n_t(L+1)}\)
Final Remarks

- Radio self-interference cancellation continues to be a challenging problem which requires significant advances in both hardware and algorithm designs - an interdisciplinary topic.

- From algorithm point of view, there are several fundamental research issues such as: Given the model

\[ p_{res} = g^T A g + g^T b + c + \text{noise} \]

- How to optimally design the training vectors of \( g \) to estimate \( A, b \) and \( c \) subject to some prior distributions?
- If we have a dynamic model for \( A, b \) and \( c \) (or a recursive algorithm for estimating them), how to design a control algorithm that concurrently adjusts \( g \) to drive \( p_{res} \) to the minimum?
100MHz bandwidth centered at 2.4GHz. The top curve is before tuning. The 3rd curve is after tuning the 1st c-tap. The 2nd curve is after tuning the 2nd c-tap. The 4th curve is after re-tuning the 1st c-tap. The bottom curve is the receiver noise.
References


