Channel Estimation

Fuyun Ling

fling@twinclouds.com

Outline

- Introduction
- Channel Estimation of Single-Tap Channels
- Channel Estimation of ISI Channels in Single Carrier Systems
- Channel Estimation in OFDM systems
- Channel Estimation for Echo/Self-Interference Cancellation
- Summary, Discussion and Concluding Remarks

fling@twinclouds.com

INTRODUCTION

fling@twinclouds.com

Communication Channels

- In digital communication systems, there are three key components: transmitter (Tx), receiver (Rx) and transmission medium (Channel)
- The Channel is the main source of impairments
- The impairments include:
  - Linear distortion
  - Non-linear distortion
  - Additive noise and interference
- One key task of receiver design and implementation is to effectively deal with channel linear distortion
- Accurate estimation of channel characteristics is the key to combat such impairments in order to achieve best possible receiver performance

fling@twinclouds.com

Classification of Communication Channels

- Transmission channels can be modeled as a linear system with one or more inputs and one or more outputs
- The basic channel type is the single input/single output (SISO) channel, or simply called channel
  - overall channel impulse response (including Tx and Rx filters) is usually longer than T, the Tx symbol interval
    - The time delay spread is due to the channel's finite bandwidth,
    - It may also be caused by multipath from reflections in radio propagation
- A communication channel also corrupts the transmitted signal with additive noise and interference
  - The additive noise is usually assumed to be white and Gaussian distributed (AWGN)
- A multiple input multiple output channel (MIMO) can be viewed as a matrix of multiple SISO channels

fling@twinclouds.com

Model of Signal Received over Channel

- The received baseband signal can be expressed as:

$$r(t) = \sum_{k=-\infty}^{\infty} a_k h(t - kT - {}^{2}) + z(t)$$

- Here we assume:
  - Perfect carrier synchronization and t < T,
  - where: T – Tx symbol interval, a_k – Tx symbol, $z(t)$ is AWGN, and h(t) – overall channel impulse response (CIR)
- In digital communication systems, the continuous received signal is sampled to generate digital samples
- The sampling interval is usually selected as T or a fraction of T, typically T/2
  - T/2 usually satisfies Nyquist criterion, thus no loss of information in the analog signal

fling@twinclouds.com

Discrete Model of Signal and Channel

- The sampled base-band signal can be expressed as:

$$r(nT)=\sum_{k=-\infty}^{\infty} a_k h\left((n-k)T - {}^{2}\right)+z(nT)$$

for T sampling. For T/2 sampling, it can be viewed as two T spaced sample streams, such that

$$r_i(nT)=\sum_{k=-\infty}^{\infty} a_k h\left((n-k)T + iT/2 - {}^{2}\right)+z_i(nT),\ i=0,1$$

  – $r_i(nT)$, i = 0, 1, are decoupled due to cycrostationality of $r(t)$
  – $z_i(nT)$, i = 0, 1, are correlated
  – The sampled CIR $h_i(mT) \cong h\left(mT + iT/2 - {}^{2}\right)$ can be viewed as with finite support, i.e., $h_i(mT) = 0$, for $m < m_1$ and $m > m_2$

- The objective of channel estimation is to estimate $h_i(mT)$ as accurately as possible

Single-Tap and ISI Channels

- Assume T – Spaced sampling, there will be interference from adjacent symbols to the symbol of interest a_n. Such a channel has inter-symbol interference (ISI).
- The channel is ISI-free if and only if $m_1 = m_2 \equiv m_0$.
  - The channel is called single-tap or single-path channel
- Single tap channel is the simplest channel and of special interest for channel estimation
- In these cases we encounter single-tap channels:
  - When the channel satisfies Nyquist Criterion
  - The channel of a Rake receiver finger after despreading in CDMA systems
  - Frequency domain channel of a subcarrier in OFDM systems

fling@twinclouds.com

Reference or Pilot based Channel Estimation

- In modern communication systems, known data symbols, called pilot, or reference, are often used
  - They are often sent together with data signal, at the beginning of transmission, during data burst, or along with data signals
- Pilot signals can greatly facilitate for channel estimation
  - Once the reception is established, if data detection has little or no error, detected data symbols can be used for channel estimation
- Optimization of pilot to traffic (P/T) ratio
  - Pilot does not carry information
    - Higher pilot power yields higher overhead
  - Higher pilot power improves channel estimation quality
    - Better channel estimates mean better receiver performance
  - Trade-off need to be made to achieve best system performance
- Even with overhead, pilot assisted coherent demodulation could still have better performance than noncoherent schemes

fling@twinclouds.com

Application of Channel Estimation

- Estimates of single tap channels can be use for correcting symbol phase rotation and weighting to generate optimal decoding metric, e.g., log-likelihood-ratio (LLR), in both single carrier and OFDM systems
- Estimates of multi-tap (ISI) channel can be used to facilitate combating ISI, i.e. equalization
  - Used directly in MLSE Equalizer
  - Can be used to compute the coefficients of linear or decision feedback equalizer (DFE)
- Channel estimates can be used to set optimal timing
- Channel estimation plays a key role in echo cancellation for wired and wireless communications

fling@twinclouds.com

ESTIMATION OF SINGLE-TAP CHANNELS

fling@twinclouds.com

Examples of Single-Tap Channel in Digital Communication Systems

- Output generated from an ISI-free channel sampled at the Tx symbol interval T with the optimal timing
  - if the overall channel frequency response satisfies Nyquist Criterion, i.e., the aliased spectrum of the channel is a constant:

$$H(\mu) = \text{const}, -1/2T \le \mu < 1/2T$$

  - In a Rake receiver in CDMA communication systems, each rake finger can be viewed as a single path channel
    - The inter-chip-interference can be viewed as a part of additive noise and they are suppressed by despreding
  - At the iFFT output of an OFDM receiver, each bin corresponding to a subcarrier can be treated a single-tap frequency domain channel

fling@twinclouds.com

Demodulation of Single-Tap Channels

- In all of these cases, the time domain channel impulse response can be expressed as a scalar $h = Ae^{jf}$
  - A and f may be time-varying (fading) or constant (static)
  - Channels always introduce additive noise/interference
- The samples at the output of the channel:

$$r(nT) = h_n a_n + z = A_n e^{j¤_n} a_n + z$$

  - a_n – Tx symbol, z additive white Gaussian noise (AWGN)
  - SNR of r(nT) (traffic) is: $¥ = |h_n|^2 E\left[|a_n|^2\right] / E\left[|z|^2\right] \cong A_n^2 P_{Ts} / ±_z^2$
- The optimum (scaled) estimate of a_n can be obtained as:

$$\tilde{a}_n = h_n^* r(nT) = |h_n|^2 a_n + h_n^* z = A_n^2 a_n + \tilde{z}$$

  - Here we assume h_n is known perfectly
  - Log-likelihood decoding metrics can be derived from $\tilde{a}_n$
  - With perfect channel coefficient, there's no SNR degradation

fling@twinclouds.com

Estimation of Single-Tap Channel Coefficient

- Practically, the channel coefficient h_n need to be estimated
- The most popular way to estimate h_n is to send known (pilot or reference) symbols, before or while sending data
- The samples corresponding to pilot: $r_{p,n} = h_n a_{p,n} + z$
- Based on one pilot symbol the scaled channel estimate can be generated as: $\tilde{h}_n = a^*_{p,n} r_{p,n} = |a_{p,n}|^2 h_n + a^*_{p,n} z$
  - $\hat{h}_n = \tilde{h}_n \big/ |a_{p,n}|^2$ is an unbiased estimate (for z to be zero mean)
  - The channel-estimate to estimation-error ratio is $\yen_{\hat{h}} \cong A_n^2 P_P \big/ \pm_z^2$
    - $P_P \cong E[|a_{p,n}|^2]$ is the pilot power
    - It is a kind like signal SNR and $\yen_{\hat{h}} = \yen_r, \; if \;\; P_T / P_P = 1$
  - $\yen_{\hat{h}}$ can be improved by averaging multiple pilot symbols

fling@twinclouds.com

Estimation of Channel Coefficient (cont.)

- Multiple pilot symbols can be used to reduce estimation error if the channel is stationary or change slowly

$$\tilde{h} = \sum_{k=0}^{N-1} a_{p,k}^{*} r_{p,k} \cong \sum_{k=0}^{N-1} \tilde{r}_{p,k} \cong N\,|a_p|^2\, h + z',\ \ \hat{h} = \tilde{h} \big/ N\,|a_p|^2$$

where $\sigma_{z'}^{2} = N\,|a_p|^2\,\sigma_z^2$, assume $|a_{p,k}| \equiv |a_p|$, as usually the case

  - It’s called *channel estimation by pilot sequence correlation*
  - Its channel estimate to error ratio, $\gamma_{\hat{h}} = N A_n^2 P_P \big/ \sigma_z^2$, i.e., N times higher than using single pilot symbol
  - Assuming the channel coefficients do not change during the estimation period and equal to h
  - The samples $r_{p,k}$ do not need to occur every T (sparse pilot)
- Channel estimates can be used for data demodulation
  - Compared to the true channel coefficient, there will be degradation, determined approximately by $\gamma / \gamma_{\hat{h}}$

fling@twinclouds.com

Optimal Static Channel Estimation

- The maximum likelihood (ML) estimator
  - We want to estimate h based on $r_{p,k}$ and $a_{p,k}$, k = 0, …, N–1
    - Define $\tilde{r}_{p,k} = a^{*}_{p,k}\tilde{r}_{p,k} = h + a^{*}_{p,k}z$, then, $E\left[\tilde{r}_{p,k}\right] = h$
    - Without loss of generality, we assume $|a_{p,k}|$=1, we have $\pm^{2}_{z'} = \pm^{2}_{z}$
    - $\tilde{r}_{p,k}$'s are complex Gaussian, the joint pdf, or likelihood function is:

$$p(\mathbf{r}\,|\,h) \cong \prod_{k=0}^{N-1} p(\tilde{r}^{(r)}_{p,k},\tilde{r}^{(i)}_{p,k}\,|\,h^{(r)},h^{(i)}) = \frac{1}{2\text{®}\pm^{2}}\prod_{k=0}^{N-1}\exp\left(-\frac{1}{2\pm^{2}}[(\tilde{r}^{(r)}_{p,k}-h^{(r)})^{2}+(\tilde{r}^{(i)}_{p,k}-h^{(i)})^{2}]\right)$$

where $\tilde{r}^{(r)}_{p,k}$, $\tilde{r}^{(i)}_{p,k}$, $h^{(r)}$ and $h^{(i)}$ are the real and imaginary parts of $\tilde{r}_{p,k}$ and h, respectively, $\pm^{2} = \pm^{2}_{z}/2$ is the real and imaginary noise variances.

  - The joint log-likelihood function is

$$\ln p(\mathbf{r}\,|\,h) = \ln\left(\prod_{k=0}^{N-1} p(\tilde{r}^{(r)}_{p,k},\tilde{r}^{(i)}_{p,k}\,|\,h^{(r)},h^{(i)})\right) = -\ln(2\text{®}\pm^{2}) - \sum_{k=0}^{N-1}\frac{(\tilde{r}^{(r)}_{p,k}-h^{(r)})^{2}+(\tilde{r}^{(i)}_{p,k}-h^{(i)})^{2}}{2\pm^{2}}$$

  - Take the derivative with respect to $h^{(r)}$ and $h^{(i)}$, we have

$$\frac{\partial}{\partial h^{(r)}}\ln p(\mathbf{r}\,|\,h) = \frac{1}{\pm^{2}}\left(\sum_{k=0}^{N-1}\tilde{r}^{(r)}_{p,k} - Nh^{(r)}\right) \text{ and } \frac{\partial}{\partial h^{(i)}}\ln p(\mathbf{r}\,|\,h) = \frac{1}{\pm^{2}}\left(\sum_{k=0}^{N-1}\tilde{r}^{(i)}_{p,k} - Nh^{(i)}\right)$$

fling@twinclouds.com

Optimal Static Channel Estimation (cont.)

- The maximum likelihood (ML) estimator (cont.)
  - The estimate is an ML estimate if both of the derivatives equal to zero and it is achieved, i.e. $\sum_{k=0}^{N-1} \tilde{r}_{p,k}^{(r,i)} = N\hat{h}^{(r,i)}$
  - Thus $\hat{h} = \sum_{k=0}^{N-1} a_{p,k}^{*} r_{p,k} \Big/ N \,|\, a_p \,|^2$ is an ML estimate of h.
  - The mean and variance of the estimation error are:

$$E\left[\frac{1}{N}\sum_{k=0}^{N-1} a_{p,k} r_{p,k}\right] = h \text{ and } |\,\hat{h} - h\,|^2 = \frac{1}{N^2} E\left[\left|\sum_{k=0}^{N-1} a_{p,k} z_k\right|^2\right] = \frac{N\pm_z^2}{N^2} = \frac{\pm_z^2}{N}$$

    - Assume that
      - Each of the noise variances of the real and imaginary parts are half of that, i.e., $\pm^2 \big/ N = \pm_z^2 \big/ 2N$
      - The cross-correlation between real and imaginary noises are zero

fling@twinclouds.com

Optimal Static Channel Estimation (cont.)

- Cramer-Rao bound of estimates
  - Definition: The variance of any *unbiased* estimator is bounded by the reciprocal of the Fisher information $\mathrm{var}(\hat{h}) \geq -1 \big/ I(\hat{h})$
    - For complex variable estimate, the Fisher information is a 2x2 matrix.
  - The elements of the Fisher information matrix for the channel estimates discussed above are

$$\mathbf{I}(\hat{h}) = \begin{pmatrix} \frac{\partial^2}{(\partial h^{(r)})^2} \ln p(\mathbf{r} \mid h) & \frac{\partial}{\partial h^{(r)} \partial h^{(i)}} \ln p(\mathbf{r} \mid h) \\ \frac{\partial}{\partial h^{(i)} \partial h^{(r)}} \ln p(\mathbf{r} \mid h) & \frac{\partial^2}{(\partial h^{(i)})^2} \ln p(\mathbf{r} \mid h) \end{pmatrix} = \begin{pmatrix} -N/\pm^2 & 0 \\ 0 & -N/\pm^2 \end{pmatrix}$$

  - Cramer-Rao bound of the estimate of h: $-\left(\mathbf{I}(\hat{h})\right)^{-1} = \begin{pmatrix} \pm^2/N & 0 \\ 0 & \pm^2/N \end{pmatrix}$
  - Thus, the ML channel estimate indeed achieves the Cramer-Rao Bound

fling@twinclouds.com

Degradation due to Channel Estimation Error

- As shown above, channel estimation errors always exist
- Using the estimated channel in demodulation of the input samples, we have:

$$\tilde{a}_n = \hat{h}_n^* r(nT) = (h_n^* + \delta_h)(h_n a_n + z) = |h_n|^2 a_n + \delta_h h_n a_n + h_n^* z + \delta_h z$$

- The last 3 terms are independent and their total variance is

$$E[|\delta_h h_n a_n|^2 + |h_n^* z|^2 + |\delta_h z|^2 = |h_n a_n|^2 E[|\delta_h|^2] + |h_n^*|^2 \sigma_z^2 + E[|\delta_h|^2]\sigma_z^2$$

- The output SNR can be expressed as:

$$\frac{\left(|h_n|^2 E[|a_n|^2]E[|\delta_h|^2] + |h_n^*|^2 \sigma_z^2 + E[|\delta_h|^2]\sigma_z^2\right)}{E[|a_n|^2 |h_n|^4} = \left\{\frac{E\left[|\delta_h|^2\right]}{|h_n|^2} + \frac{\sigma_z^2}{E[|a_n|^2 |h_n|^2} + \frac{E[|\delta_h|^2}{|h_n|^2}\frac{\sigma_z^2}{E[|a_n|^2 |h_n|^2}\right\}^{-1}$$

$= \left\{\frac{1}{\gamma_h} + \frac{1}{\gamma} + \frac{1}{\gamma_h}\frac{1}{\gamma}\right\}^{-1}$ – where γ is SNR and $1/\gamma_h$ is the relative channel estimation error

  - If $\gamma_h >> 1$ and $\gamma >> 1$, we can ignore the second order term (the last term) and the output SNR is approximately $\frac{\gamma}{1+\gamma/\gamma_h}$

fling@twinclouds.com

Single-Tap Fading Channel Estimation

- In wireless communications, the channel is usually not static but time varying, i.e., fading
- The fading channel coefficient $h(t)$ can be viewed as a low-pass random process with a Doppler power spectrum ¤$_d(f)$
- Fading channel coefficients are usually estimated by using pilot signal send together with traffic data signal
  - Pilot signal can be mixed with Traffic data in CDM (cdma system), TDM (for different types of systems) and FDM (for OFDM systems).
  - At the receiver, pilot samples are generated every T_p (may be different from data symbol interval T) with or without spreading
  - Pilot samples can be expressed by $r_{p,k} = h(kT_p)+z_k$
  - z_k is the additive noise with the power spectrum of ¤$_n(f)$

fling@twinclouds.com

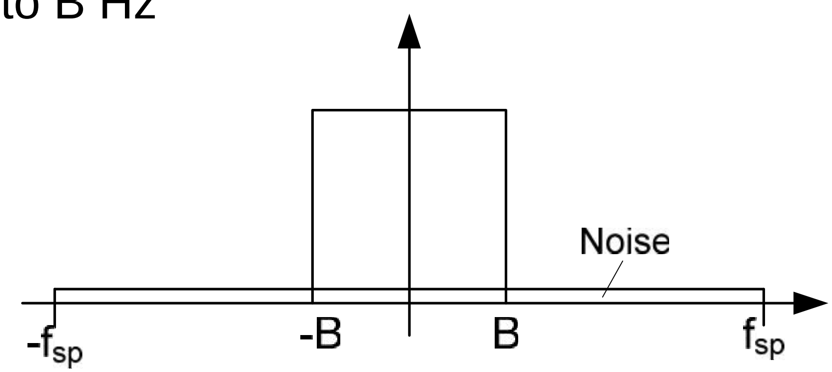
Optimal Fading Channel Estimation

- For static channel, the channel estimation accuracy is proportional to the integration time
- For fading channel the integration time cannot be too long
  - Long integration time that exceed channel coherent time will degrade channel estimate
- Optimal single-tap channel estimator is a linear Wiener filter if the power spectrum of the channel fading is known
  - The Wiener filter frequency response is $¤_d(f)/[¤_d(f)+¤_n(f)]$
- When the Doppler spectrum is not known or changes with time, a low pass rectangular filter with a bandwidth B equal to the maximum Doppler frequency can be used
  - Such a filter is an optimal *Robust Estimator*. i.e., it yields the best performance under the worst channel condition.

fling@twinclouds.com

Optimal Fading Channel Estimation (cont.)

- It has been shown that the complex conjugate of channel estimates multiplied by received signal sample is optimal for data symbol demodulation in a class of popular channels
  - This class of channels include static, Rayleigh and Nakagami fading channels
  - The optimality means the decoding metric such generated is ML
- Performance analysis of the optimal robust channel estimator
  - The optimal robust estimator is a linear filter with rectangular pass-band frequency response from –B to B Hz
  - Denote the pilot symbol sampling rate by $f_{s,p}$, the ratio of the channel energy to channel estimation error, will be $f_{s,p}/2B$ time higher than SNR of the pilot symbol



fling@twinclouds.com

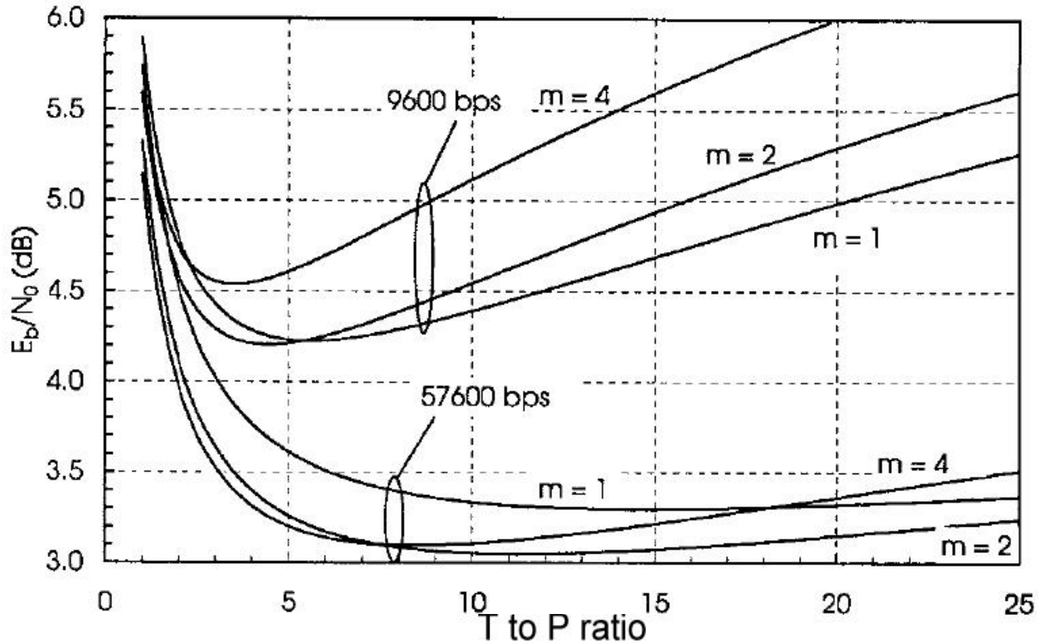
Performance Optimization of System with Pilot Based Coherent Demodulation

- In such a system, the total power is split between traffic (data) and pilot channels (P_T and P_p)
  - The higher pilot power, the higher overhead – worse system performance
    - Data SNR γ is proportional to $P_T/(P_T+P_p) = \alpha$
  - Higher pilot power yields better channel estimate, i.e., γ_h, yield better system performance
    - γ_h is proportional to $P_p/(P_T+P_p) = 1 - \alpha$
  - SNR at demodulator output is equal to $\left\{\frac{1}{\gamma_h}+\frac{1}{\gamma}+\frac{1}{\gamma_h}\frac{1}{\gamma}\right\}^{-1}$ has a maximum at certain α
- Traffic to pilot ratio is an important design factor for 3G and 4G wireless systems and widely adopted in many standard specifications

fling@twinclouds.com

Performance Optimization of System with Pilot Based Coherent Demodulation (cont.)

- Example:



  - Required E_b/I_0 versus T to P ratio at cut-off rate R_0= 0.25 in a CDMA system design for Rayleigh and Nakagami 2 and 4
    - Information bit rate of 9600 b/s and 57600 b/s
    - Maximum Doppler frequency B = 300Hz

(IEEE Transections on Communications, VOL. 47, NO. 10, pp. 1583-1592, October 1999)

fling@twinclouds.com

ESTIMATION OF MULTIPLE TAP CHANNELS

fling@twinclouds.com

General Description of Multiple Tap Channels

- In general, communication channels are not ISI free
  - Such channels can be modeled as FIR filters with multiple taps
  - Even if a channel satisfies Nyquist criterion, it is ISI free only when samples at the right timing phase
  - The single tap channel is a special case of a multi-tap channel.
  - However, some of the results for single tap channel case can be extended to the multi-tap channel case
- T-Spaced signal samples from multi-tap channels:

$$r_n \cong r(nT) = \sum_{k=-K_1}^{K_2} a_{n-k} h\left(kT - {}^{2}\right) \cong \sum_{k=-L_1}^{L_2} a_{n-k} h_k$$

  - h(t) is the continuous-time channel impulse response (CIR)
- T/2-Spaced signal samples from multi-tap channels:

$$r_{n,i} \cong r(nT + iT/2) = \sum_{k=-K_1}^{K_2} a_{n-k} h\left(kT - {}^{2} - \frac{iT}{2}\right) \cong \sum_{k=-K_1}^{K_2} a_{n-k} h_{k,i}, \;\; i = 0,1$$

fling@twinclouds.com

Estimation of Multiple Tap Channels

- Multi-tap channels are usually estimated by using pilots
  - Pilots are sent at the beginning of data transmission, embedded within a data burst, or interleaved with data
- Pilot assisted T-Spaced multi-tap channel estimation:
  - The received signals corresponding to pilot are:

$$r_{p,n} \cong r_p(nT) = \sum_{k=-K}^{K_2} a_{p,n-k} h_k + z_n$$

  - The correlation based estimate of h_k can be expressed as

$$\hat{h}_k = \sum_{n=N_1}^{N_2} a^*_{p,n-k} r_{p,n} = \sum_{n=N_1}^{N_2} \sum_{l=-K_1}^{K_2} a^*_{p,n-k} a_{p,n-l} h_k + z' = \sum_{n=N_1}^{N_2} \left| a_{p,n-k} \right|^2 h_k + \sum_{\substack{l=-K_1 \\ l \neq k}}^{K_2} h_k \sum_{n=N_1}^{N_2} a^*_{p,n-k} a_{p,n-l} + z'$$

  - This is similar to the correlation based single-tap channel estimation with an additional error term
    - This error term can be ignored or eliminated if
      - Long correlated sequence summation and large noise (CDMA case)
      - or, the sequence $\{a_{p,n}\}$ has the property $\sum_{n=N_1}^{N_2} a^*_{p,n-k} a_{p,n-l} = 0, \text{ for } n \neq l$
  - CE error is proportional to the number of taps to be estimated

fling@twinclouds.com

Sequence Examples for Channel Estimation

- Sequences in CDMA systems
  - These sequences are uncorrelated but the practically used short correlations are not necessary equal to zero for $k \neq l$
  - Reduces noise/interference proportional to integration length
- GSM midamble sequence
  - Sequence length = 22
  - Correlation length = 16 by design
    - Reduce noise variance by a factor of 16, i.e., 12 dB
  - The autocorrelation is equal to zero for $1 \leq |k - l| \leq 6$
    - The maximum channel estimate length = 6
  - It is possible to use least-squires estimation to increase the channel estimate length and/or improve noise reduction
    - E.g. with the same noise reduction factor, the estimation length increases to 12

fling@twinclouds.com

Sequence Examples for CE in OFDM

- Frequency domain property of sequence with impulsive circular correlation
  - If a sequence of length N has an impulsive circular correlation i.e. $$\sum_{n=0}^{N-1} x^*[n]x[(n+l)_{Mod\ N}] = \begin{cases} NE[|x(n)|^2] & l=0 \\ 0 & l \neq 0 \end{cases}$$
  - Its DFT outputs have the same magnitude, i.e., $|X(k)| = C \quad \forall k$
  - Two segments of such sequence cascaded can be used for channel estimation up to length N
- An OFDM symbol with one non-zero value every other M-1 zero values and constant magnitude
  - The time domain sequence is periodic with M periods
  - The time domain sequences have impulsive circular correlation
  - Can estimate a channel of N_{FFT}/M samples long

fling@twinclouds.com

Sequence Examples for CE in OFDM (cont.)

- Zadoff-Chu Sequences in LTE
  - Zadoff-Chu sequence ("chirp" sequence)

$$ZC_{N,M}[k] = \begin{cases} \exp(-j\,M®k^2/N) & \textit{for } N \textit{ even} \\ \exp(-j\,M®k(k+1)/N) & \textit{for } N \textit{ odd} \end{cases}$$

    - M and N are relative prime
    - They have impulsive circular autocorrelations
    - They have constant magnitudes in time domain (low peak to average)
      - Better in this regard than the previous general results
  - Zadoff-Chu sequences used in LTE as Primary Synchronization Signal (PSS)
    - Length 63 with M = 35, 29, or 34 (3 cell-group indices)
    - The 31's value (DC) forced to zero
      - Autocorrelation no longer perfect but still pretty good

fling@twinclouds.com

Sub-T Spaced Channel Estimation

- T-Spaced channel estimates do not satisfy Nyquist criterion
- It is sufficient to use T/2 spaced channel estimates
- T/2 Spaced channel estimate is essentially two sets of T-spaced channel estimates due to the cyclo-stationality of signals
- We can apply the above estimation methods twice on to the signal samples $r(nT)$ and $r(nT+T/2)$
- If it is desirable to perform T-spaced processing, e.g. equalization, the channel estimate at certain offset with higher energy indicates the signal sample sequence at this offset should be used

fling@twinclouds.com

Other Channel Estimation Methods

- When Tx data symbols are recovered (known), they can be used for channel estimation as well
  - The Tx symbols can be the same as the data to be received or from other channels
  - Once the data recovery and channel estimation can be performed iteratively ("turbo channel estimation")
  - It is popular for wireline communications where known symbols are only inserted at the beginning of transmission
- Multi-tap channels can also be estimated using LMS or RLS adaptive algorithms
  - The sequence for estimation can be pilot sequence or recovered data sequence
  - The details of such estimation algorithm will be discussed later for echo/self-interference cancellation

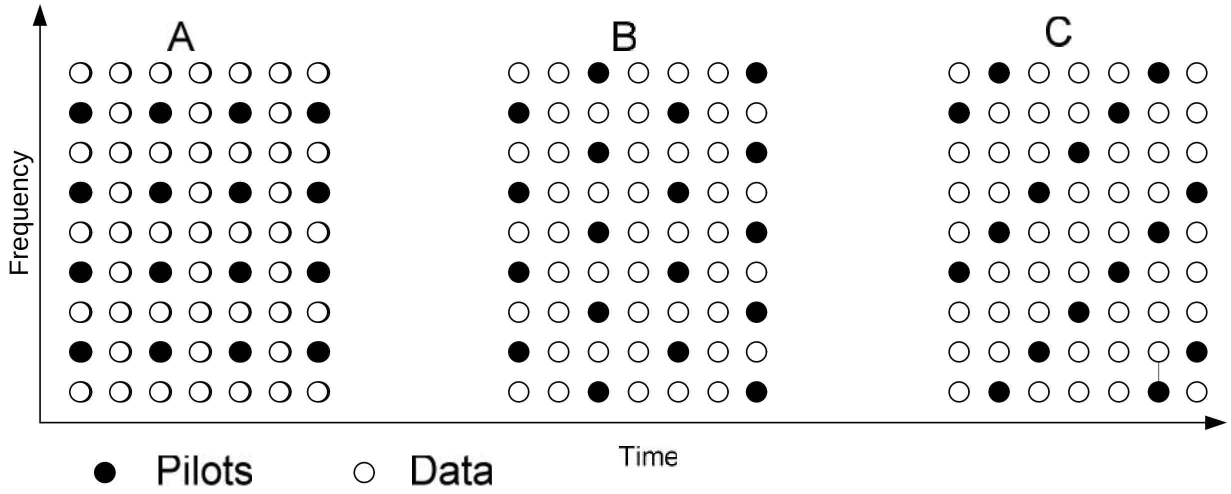
fling@twinclouds.com

CHANNEL ESTIMATION IN OFDM SYSTEMS

fling@twinclouds.com

Frequency Domain Channel Estimation

- Modern OFDM systems almost always use pilot based channel estimation
  - Coherent demodulation has better performance than non-coherent, e.g., differential, demodulation, even with pilot overhead
  - Frequency domain (FDM) pilots based CE is most popular
    - With proper cyclic prefix (CP) there’s no inter-chip interference (ICI)
  - Typical FDM pilot patterns:



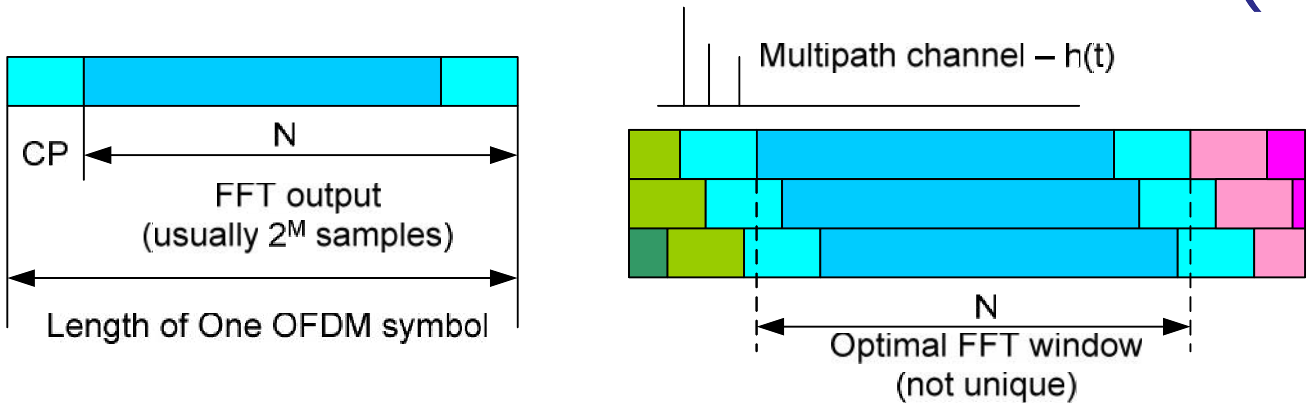
fling@twinclouds.com

Transmission and Demodulation in OFDM

- In an OFDM system, data are organized as vectors of N modulation symbols (OFDM symbols) for transmission
- Such OFDM symbols are converted to time domain by iFFT and prepended cyclic prefixes (CP) for transmission
- During transmission the time domain symbols with CP are convolved with channel
- Each N-samples long time domain sequence of one OFDM symbol with CP removed is converted to frequency domain (FD) by FFT
  - If the channel span is less than CP, it has one complete period of the data circularly convolved with channel and no ISI
- The resulting FD OFDM symbol is the Tx OFDM symbol multiplied by channel frequency response

fling@twinclouds.com

Transmission and Demod. in OFDM (cont.)



Transmitted signal

Received signal

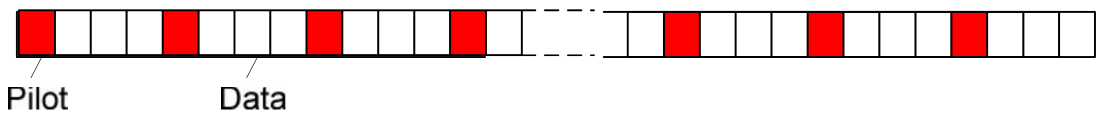
Time Domain

- Frequency domain data vectors:
  - Transmitted (FD) data vector: $[X(0), X(1), ..., X(k), ...X(N-1)]$
  - Received (FD) data vector at the FFT output: $[Y(0), Y(1), ..., Y(N-1)]$
    - where Y(k) = H(k)X(k) and {H(k)} are the FFT of sampled {h(t)}
- Data symbols $\{X(k)\}$ can be recovered by multiplying the the conjugate of the estimated FD channel value $\hat{H}(k)$
  - $\hat{H}(k)$ is the true value of $H(k)$ plus and error term due to noise
  - $\{\hat{H}(k)\}$ can be obtained from FFT of the estimated CIR $\hat{h}(n)$

fling@twinclouds.com

Frequency Domain CE with FDM Pilots

- An example of an OFDM symbol with FDM pilots



- Assume there are one pilot in every M subcarrier, say the symbols on mM^th^ subcarriers are pilot, the estimates of the mM^th^ subcarriers are $\hat{H}(mM) = X^*(mM)Y(mM) / |X(mM)|^2$
- The objective is to compute the estimates where the data subcarriers located
- An N/M points inverse FFT of $\{\hat{H}(mM)\}$ yields an estimate of channel CIR, as long as the channel span is less than N/M
- An N point FFT of N(M-1)-zero padded at the end of the CIR estimate yields the FD channel estimate of all N subcarriers
- This is equivalent to interpolate $\{\hat{H}(mM)\}$ by a filter with magnitude response of $\sin(k\circledR M/N) / \sin(k\circledR/N)$

fling@twinclouds.com

Frequency Domain CE with FDM Pilots (cont.)

- Other methods used for FD CE interpolation
  - Essentially, any interpolation methods can be used
  - Commonly used methods are low-pass FIR filtering, linear interpolation with different order, and non-linear interpolation, e.g., spline based methods
  - The effect may be different, the objective is to achieve the best performance/complexity ratio
- Performance impact of channel estimation error
  - The pilot and data symbols have the same noise variance
  - Their SNRs, denoted as γ_d and γ_p depend on their symbols' powers
  - The SNR of the demodulated symbol can be expressed by:

$$\gamma_s = \left(\gamma_d^{-1} + \gamma_p^{-1} + \gamma_d^{-1}\gamma_p^{-1}\right)^{-1}, \text{ or } \approx \left(\gamma_d^{-1} + \gamma_p^{-1}\right)^{-1} \text{ for } \gamma_d >> 1, \gamma_p >> 1$$

  - If $\gamma_d = \gamma_p$ the degradation will be greater than 3 dB

fling@twinclouds.com

Frequency Domain CE with FDM Pilots (cont.)

- Improving quality of channel estimation (QCE)
  - Direct interpolation of frequency domain channel estimates from each OFDM symbol yields undesirable degradation
  - The following methods can be used to achieve QCE improvements

  (1) Better averaging in frequency domain (FD)
    - Interpolation can be improved if we have an idea of the statistics of CIR
      - Instead of using rectangular window, we can use a window fit the pdf of CIR
        - » This is the so-called MMSE channel estimation (may not be practical)
    - Elimination of noisy time domain channel (CIR) taps
      - If the actually taps are less than M – the maximum taps can be estimated

      A. Generating the CIR from the FD CE with iFFT

      B. Delete the taps likely taps due to noise

      C. Convert the remaining taps back to FD with FFT
        - » If only L taps are used to generate the N FD CE by the FFT, the noise variance is reduced by a factor of M/L
        - » i.e., γ_p is improved by a factor of M/L

fling@twinclouds.com

Frequency Domain CE with FDM Pilots (cont.)

- Improving channel estimation accuracy (QCE) (cont.)

  (2) Averaging in time domain

  - For channel coherent time longer than the time between the OFDM symbols with FDM pilots, QCE can be improved by averaging
  - If the FDM pilots are aligned in frequency in different OFDM symbols (Pattern A), the averaging can be done directly between FD CEs
  - If they are not aligned (Pattern B and C), averaging can be done in time domain by first converting FE CDs to CIRs by iFFT
  - The averaging window need to selected care to reduce both noise variance and delay
- Timing synchronization using FD channel estimate
  - The objective of OFDM timing synchronization is to determine FFT window placement
  - Frequency domain CE are converted to time domain by iFFT
  - The resulting CIR can be used for time synchronization

fling@twinclouds.com

Impact of Guardband to FD CE

- The iFFT/FFT method with CIR Clean-up is close to optimal for interpolation if the condition of circular convolution is satisfied
  – This is true with sufficiently long CP and no guardbands
  – With guardbands, it is still close to optimal in middle of the OFDM symbol.
- At the edge of OFDM symbol, it is no-longer optimal because the pilots in guardbands are missing
  – The FD CEs are always degraded for any interpolation method
  – The best performance can be achieved if using optimally weighted interpolation at the edges but using the iFFT/FFT in the middle
  – The performance improvement is relatively limited
    - Approximately 0.1-0.3 dB overall improvement may be achieved

fling@twinclouds.com

CHANNEL ESTIMATION FOR ECHO/SELF-INTERFERENCE CANCELLATION

fling@twinclouds.com

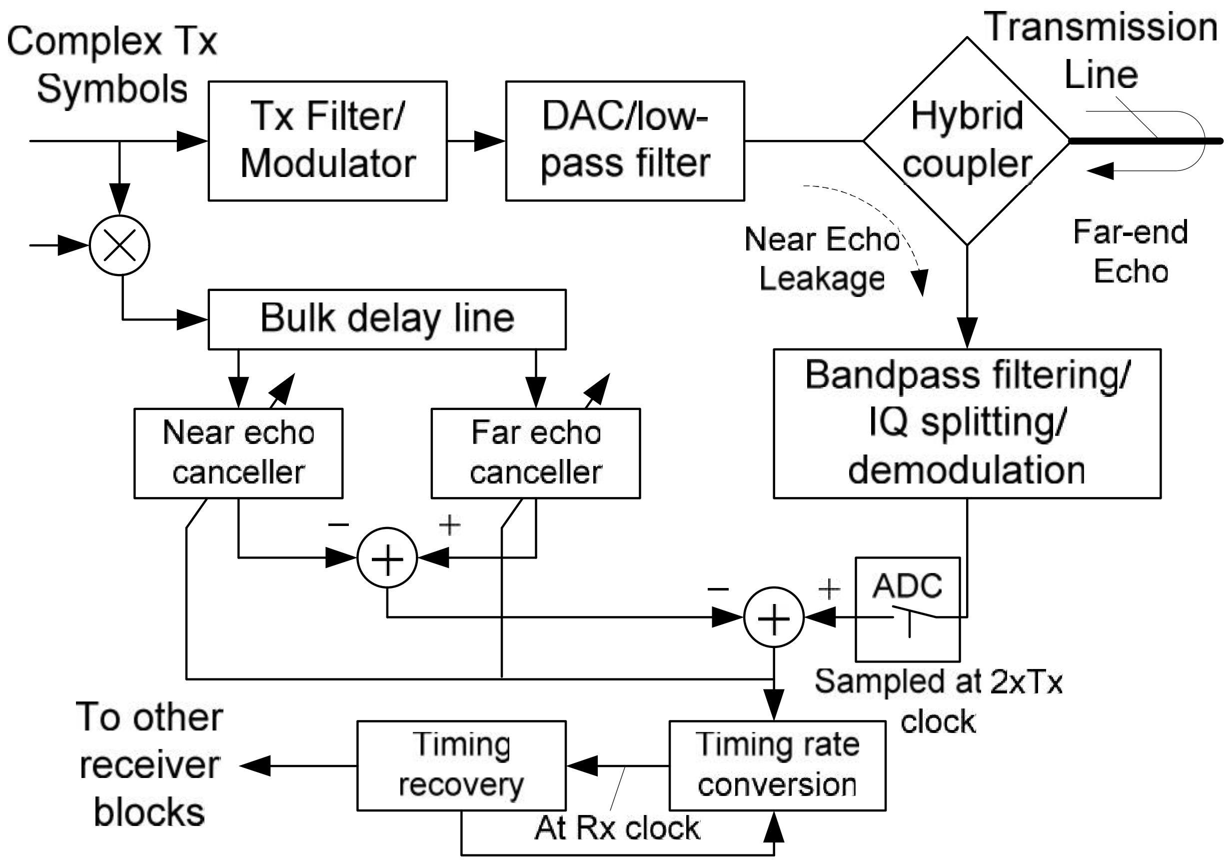
Echo/Self-Interference Cancellation

- Both technologies achieve full-duplex communication using the same frequency band and at the same time by
  - Creating isolation of Tx signal from Rx input
  - Using known Tx data and accurately estimated interference channel to synthesize a replica of the interference to be subtracted from the received signal
- Echo-Cancellation for wireline communication products has been known since 60's and utilized in commercial products since mid 80's
- Self-Interference cancellation for wireless communications has attracted attention in recent years
- Both are based on very similar operation principles

fling@twinclouds.com

Echo/Self-Interference Cancellation (cont.)

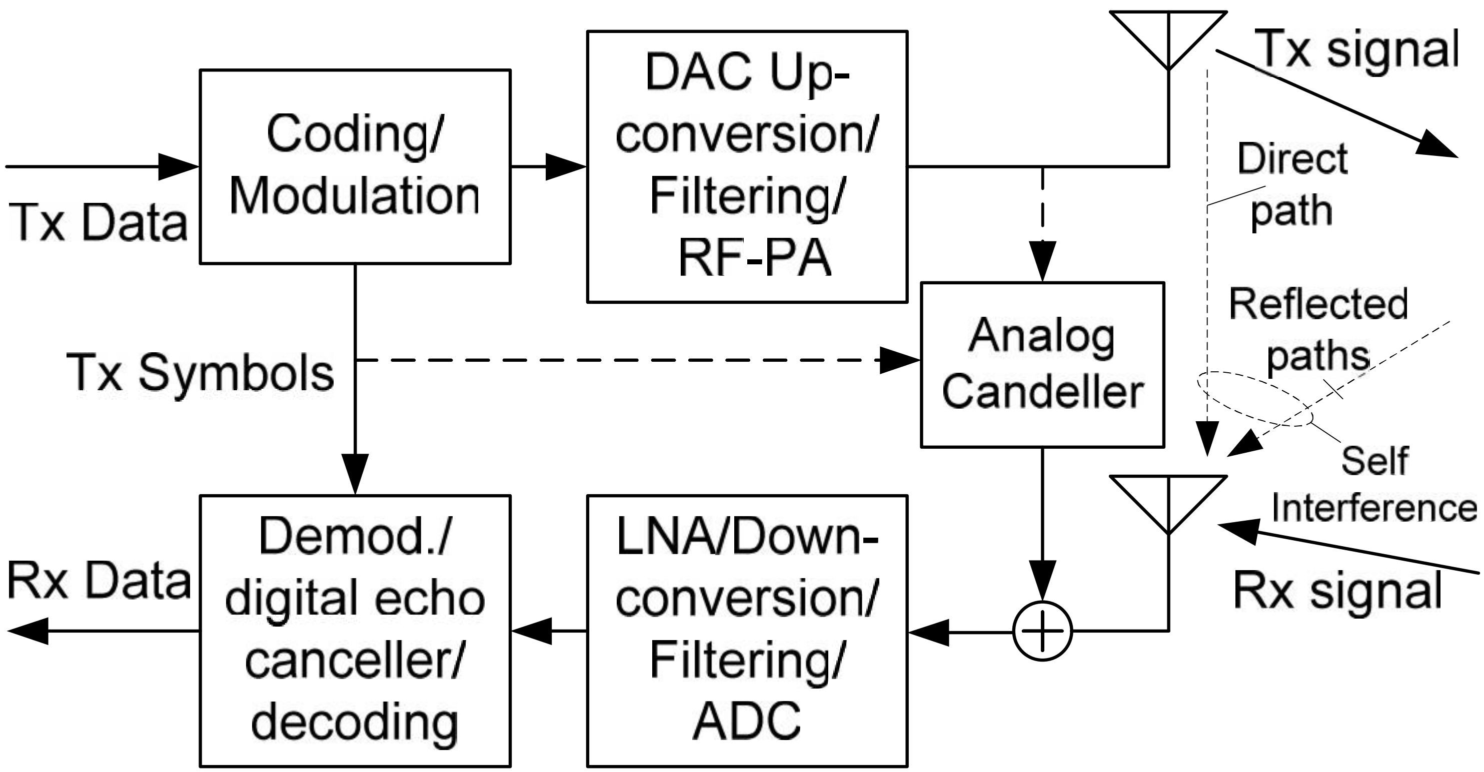
- Full-duplex wireline modem utilizing echo-cancellation



fling@twinclouds.com

Echo/Self-Interference Cancellation (cont.)

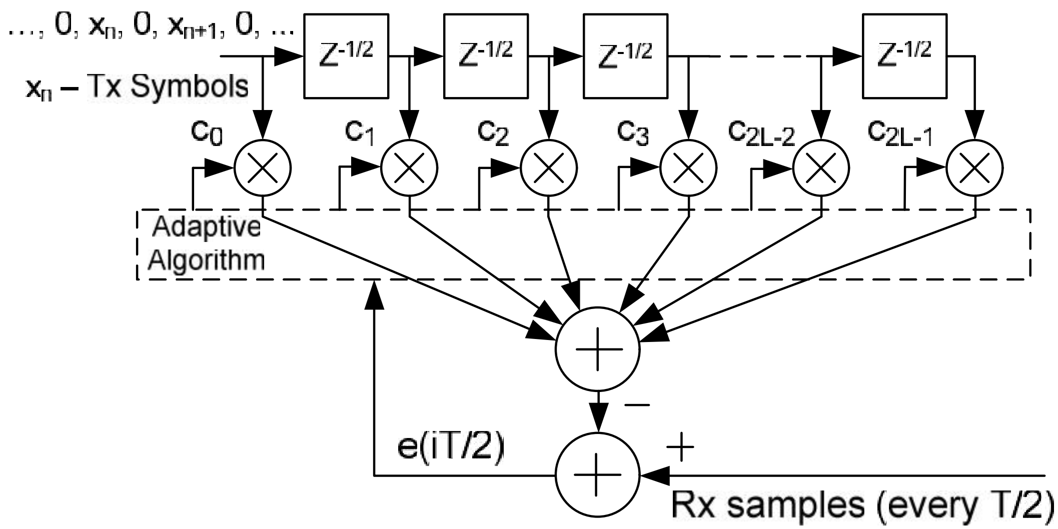
- In-band full-duplex wireless system based on self-interference cancellation



fling@twinclouds.com

Channel Estimation for Echo/Self-Interference Cancellation

- Accurate channel estimation is the key to achieve maximum interference cancellation
- LMS algorithm is commonly used for adaptation
  - It is simple (less computation demanding) to implement
  - With uncorrelated Tx symbols as reference input, LMS can achieve the same tracking ability as exponentially weighted LS algorithm



fling@twinclouds.com

Analysis of Channel Estimation Errors

- We shall analyze two types of LMS channel estimation errors that limit the performance of the cancellers
  1. Excess mean-square-error (MSE), also called self-noise
  2. Tracking error due to interference channel variation
- The basic equations of the LMS adaptive subfilter are:
  - The estimated interference: $\hat{I}(n) = \sum_{k=0}^{L-1} x_{n-k} c_k(n)$
  - The estimation error: $e(n) = r(n) - \hat{I}(n)$
  - Coefficient update equation: $c_k(n+1) = c_k(n) + \Delta x_{n-k} e_i(n)$
- Excess MSE: $\text{£}_{ex} = (\text{«} / 2) L\text{£}$

  Where $\text{«} = \Delta / LE[|x_n|^2]$ is the normalized adaptation step size,

  L – number of coefficients, e – MSE of the irreducible error,

  – In the case of the cancellers, e – is the total received signal power

fling@twinclouds.com

Analysis of Channel Estimation Errors (cont.)

- Excess MSE analysis
  - For echo/self-interference cancellers the received signal is the "noise".
  - The higher the required SNR, the higher "noise" to the canceller
  - The residual error after cancellation should be at least 6 dB lower than the required noise level for the required Rx SNR (γ)
  - The (normalized) step-size is determined by the required receiver SNR and the length of the canceller:

$$« \le 2 \times 10^{-(¥+6)/10} \Big/ L$$

  - Examples:
    - For g = 27 dB and L = 100: m $\le 10^{-5}$
    - For g = 21 dB and L = 40: m $\le 10^{-4}$
  - Such an adaptive filter converges according to: $(1\text{-m})^n$
    - Its time constant is approximately T/(1-m)

fling@twinclouds.com

Analysis of Channel Estimation Errors (cont.)

- Tracking error analysis
  - The reference symbols input to the cancellers are uncorrelated
  - The all the modes of the canceller converges uniformly
  - The system can be modeled as a linear system with an impulse response of $h(n)=U(n)m(1\text{-}m)^n$
  - The frequency domain response is

$$H(e^{j\mu}) = \sum_{n=0}^{\infty} « e^{-j\mu n}(1-«)^n = \frac{«}{1-e^{-j\mu}(1-«)}$$

  - The difference between the LMS system frequency response and the ideal all-pass filter's response, which is a constant 1, is

$$1-H(e^{j\mu}) = 1-\frac{«}{1-e^{-j\mu}(1-«)} = \frac{(1-e^{-j\mu})(1-«)}{1-e^{-j\mu}(1-«)}$$

  - For static system, the differences is zero, i.e., the estimator is ideal
  - For time variant channel, the estimation error is not zero – it imposes an limit or the maximum achievable cancellation

fling@twinclouds.com

Analysis of Channel Estimation Errors (cont.)

- Tracking error analysis (cont.)
  - Assuming channel varies as a complex sinusoidal $e^{j\mu_0}$

$$1 - H(e^{j\mu_0}) = \frac{2(1-\text{«})(1-\cos\mu_0)}{(2-\text{«})(1-\cos\mu_0) - j\text{«}\sin\mu_0} \approx \frac{(1-\text{«})^2\mu_0^2}{(2-\text{«})^2\mu_0^2/4+\text{«}^2}$$

    - Follows square rule of ω_0 and $1/\mu$, i.e., 20 dB change in estimation error due to the change of a factor 10 in ω_0 or $1/\mu$

Limit of Cancellation	f_DT (ratio of fading frequency to Tx symbol rate)			
	10^{-7}	10^{-8}	10^{-9}	10^{-10}
$\mu = 10^{-4}$	44 dB	64 dB	84 dB	104 dB
$\mu = 10^{-5}$	24 dB	44 dB	64 dB	84 dB

  - Remarks:
    - Not going down very fast as fading frequency reduces
    - Cannot be ignored for high cancellation requirement
    - The same conclusion also applies to exponential weighting LS estimator
    - More serious for the high demanding channel estimation of echo/self-interference cancellers than for receiver channel estimation
    - The analysis based on system equation is not valid if the data are correlated

fling@twinclouds.com

SUMMARY AND CONCLUDING REMARKS

fling@twinclouds.com

Summary and Concluding Remarks

- To achieve best possible system performance, receivers must have knowledge of the communication channel
  - Direct or indirect channel estimation is essential
- Due to the existence of noise, interference, distortion and channel variation, the channel estimates can be expressed as the true channel with estimation error.
- The objective is to attain the best possible estimate under realistic channel conditions.
- The accuracies of the channel estimates are always limited by theoretical bounds such as Cramer-Rao bound
  - Practically achievable system performance must take into account of such theoretical limits

fling@twinclouds.com

Summary and Concluding Remarks (cont.)

- In this presentation we discussed:
  - Channel modeling;
  - Estimation of single-tap and multi-tap channels;
  - Optimal channel estimation under static and fading conditions
- Examples of channel estimations are given for
  - Single carrier systems, OFDM systems and Echo/self-interference cancellation
- Using detected data symbol to improve estimation accuracy
  - When considering achievable channel estimation accuracy, we mainly assume the estimation are based on known data symbols such as pilot symbols
  - By using detected data symbols, estimation accuracy can be further improved, e.g., with “Turbo” channel estimation

fling@twinclouds.com

Acknowledgement

The materials in this presentation are mostly not new. I just pulled them from the literature and from my experience accumulated during the past 30 years. Even most of these experience were also learned from others as well.

Thus, hereby I would like to thank all the researchers and engineers who contributed to this technical topic. The credits should go to the respective contributors. However, because there are so many of them, I really cannot mention all. Of cause, I should responsible for any mistake in the presentation and would sincerely appreciate if anyone would take time to point them out to me.

Sincerely,

Fuyun Ling

fling@twinclouds.com

THANK YOU!

fling@twinclouds.com