BAYESIAN SHORT-TIME SPECTRAL AMPLITUDE ESTIMATORS FOR SINGLE-CHANNEL SPEECH ENHANCEMENT

Eric Plourde

Department of Electrical and Computer Engineering
Telecommunications & Signal Processing Laboratory
McGill University

Ph.D. Defense
October 14th, 2009
OUTLINE

1 INTRODUCTION

2 BACKGROUND
   Speech enhancement
   Bayesian STSA estimation for speech enhancement

3 THESIS CONTRIBUTIONS
   1 - Further analysis of the $\beta$-SA estimator
   2 - Weighted $\beta$-SA with auditory-based parameter values
   3 - Analytical generalization of Bayesian STSA estimators
   4 - Multi-dimensional estimators allowing correlated frequency components

4 CONCLUSION
1 **INTRODUCTION**

2 **BACKGROUND**

   Speech enhancement
   Bayesian STSA estimation for speech enhancement

3 **THESIS CONTRIBUTIONS**

   1 - Further analysis of the $\beta$-SA estimator
   2 - Weighted $\beta$-SA with auditory-based parameter values
   3 - Analytical generalization of Bayesian STSA estimators
   4 - Multi-dimensional estimators allowing correlated frequency components

4 **CONCLUSION**
OUTLINE

1 INTRODUCTION

2 BACKGROUND
   Speech enhancement
   Bayesian STSA estimation for speech enhancement

3 THESIS CONTRIBUTIONS
   1 - Further analysis of the $\beta$-SA estimator
   2 - Weighted $\beta$-SA with auditory-based parameter values
   3 - Analytical generalization of Bayesian STSA estimators
   4 - Multi-dimensional estimators allowing correlated frequency components

4 CONCLUSION
OUTLINE

1. INTRODUCTION

2. BACKGROUND
   Speech enhancement
   Bayesian STSA estimation for speech enhancement

3. THESIS CONTRIBUTIONS
   1. Further analysis of the $\beta$-SA estimator
   2. Weighted $\beta$-SA with auditory-based parameter values
   3. Analytical generalization of Bayesian STSA estimators
   4. Multi-dimensional estimators allowing correlated frequency components

4. CONCLUSION
# Outline

1. **Introduction**

2. **Background**
   - Speech enhancement
   - Bayesian STSA estimation for speech enhancement

3. **Thesis Contributions**
   1. Further analysis of the \( \beta \)-SA estimator
   2. Weighted \( \beta \)-SA with auditory-based parameter values
   3. Analytical generalization of Bayesian STSA estimators
   4. Multi-dimensional estimators allowing correlated frequency components

4. **Conclusion**
WHY SPEECH ENHANCEMENT

A speech signal can be corrupted by different additive noises:

- someone else speaking,
- a car passing by,
- background music, etc.

In many speech applications, there is an advantage to remove that noise:

- Hearing aids
- Automatic speech recognition
- Speech coders
WHY SPEECH ENHANCEMENT

A speech signal can be corrupted by different additive noises:

- someone else speaking,
- a car passing by,
- background music, etc.

In many speech applications, there is an advantage to remove that noise:

- Hearing aids
- Automatic speech recognition
- Speech coders
A speech signal can be corrupted by different additive noises:

- someone else speaking,
- a car passing by,
- background music, etc.

In many speech applications, there is an advantage to remove that noise:

- Hearing aids
- Automatic speech recognition
- Speech coders
A speech signal can be corrupted by different additive noises:

- someone else speaking,
- a car passing by,
- background music, etc.

In many speech applications, there is an advantage to remove that noise:

- Hearing aids
- Automatic speech recognition
- Speech coders
Why Speech Enhancement

A speech signal can be corrupted by different additive noises:

- someone else speaking,
- a car passing by,
- background music, etc.

In many speech applications, there is an advantage to remove that noise:

- Hearing aids
- Automatic speech recognition
- Speech coders

How? ⇒ Noise reduction / speech enhancement algorithms
MAIN EXISTING SPEECH ENHANCEMENT APPROACHES

- **Spectral subtraction**
  (e.g. Boll, 1979; Virag, 1999; Hasan *et al.*, 2004)

- **Kalman filter based**
  (e.g. Paliwal *et al.*, 1987; Gannot *et al.*, 1998; Ma *et al.* 2006)

- **Subspace**
  (e.g. Ephraim *et al.*, 1995; Jabloun *et al.*, 2003; You *et al.*, 2005)

- **Bayesian STSA estimators**
  (e.g. Ephraim *et al.*, 1984; Loizou, 2005; You *et al.*, 2005; Erkelens *et al.*, 2007)

  - Perform generally better than many other approaches (Hu and Loizou, 2007).
  - Computational demand relatively modest compared to subspace or Kalman filter based approaches.
MAIN EXISTING SPEECH ENHANCEMENT APPROACHES

- **Spectral subtraction**  
  (e.g. Boll, 1979; Virag, 1999; Hasan et al., 2004)

- **Kalman filter based**  
  (e.g. Paliwal et al., 1987; Gannot et al., 1998; Ma et al. 2006)

- **Subspace**  
  (e.g. Ephraim et al., 1995; Jabloun et al., 2003; You et al., 2005)

- **Bayesian STSA estimators**  
  (e.g. Ephraim et al., 1984; Loizou, 2005; You et al., 2005; Erkelens et al., 2007)
  - Perform generally better than many other approaches (Hu and Loizou, 2007).
  - Computational demand relatively modest compared to subspace or Kalman filter based approaches.
SUMMARY OF CONTRIBUTIONS

1 - Study the negative values of $\beta$ in the $\beta$-SA estimator. (Chapter 4)

2 - Develop a new Bayesian STSA estimator ($W_\beta$-SA) combining the power law and weighting factor of existing estimators and propose values for its parameters based on auditory considerations. (Chapter 5)

3 - Develop a new general family of Bayesian STSA estimators (GWSA). (Chapter 6)

4 - Propose a new family of Bayesian STSA estimators that assumes correlated frequency components. (Chapter 7)
OUTLINE

1. INTRODUCTION

2. BACKGROUND
   Speech enhancement
   Bayesian STSA estimation for speech enhancement

3. THESIS CONTRIBUTIONS
   1 - Further analysis of the $\beta$-SA estimator
   2 - Weighted $\beta$-SA with auditory-based parameter values
   3 - Analytical generalization of Bayesian STSA estimators
   4 - Multi-dimensional estimators allowing correlated frequency components

4. CONCLUSION
SPEECH ENHANCEMENT IN STFT/STSA DOMAIN

Noisy speech: 
\[ Y_k = |Y_k| e^{j \angle Y_k} \]

Clean speech: 
\[ X_k = |X_k| e^{j \angle X_k} \]

Noise: 
\[ W_k = |W_k| e^{j \angle W_k} \]
**SPEECH ENHANCEMENT IN STFT/STSA DOMAIN**

Noisy speech: $Y_k = X_k + W_k$

Clean speech: $X_k = X_k e^{j\angle X_k}$

Noise: $W_k = |W_k| e^{j\angle W_k}$

You know $Y_k$ and you want to estimate $X_k$. 
**SPEECH ENHANCEMENT IN STFT/STSA DOMAIN**

Since phase is less perceptually significant than amplitude:

- Estimate $X_k \Rightarrow$ Short time spectral amplitude (STSA),
- Combine with the noisy speech phase $\angle Y_k$ to form estimator:

$$\hat{X}_k = \hat{X}_k e^{j \angle Y_k}$$
Bayesian STSA estimation

Bayes estimator of $\mathcal{X}_k$

$$\hat{\mathcal{X}}_k^o = \arg \min_{\hat{\mathcal{X}}_k} E\{ C(\mathcal{X}_k, \hat{\mathcal{X}}_k) \}$$

⇒ Specify appropriate statistical models (usually Gaussian i.i.d.) and a cost function $C(\mathcal{X}_k, \hat{\mathcal{X}}_k)$.

Solutions are generally provided in the form:

$$\hat{\mathcal{X}}_k^o = G_k |Y_k|$$

where $G_k$ is the gain of the estimator.
\textbf{Bayesian STSA estimation}

\[
\hat{x}_k^o = \arg\min_{\hat{x}_k} E\{C(x_k, \hat{x}_k)\}
\]

⇒ Specify appropriate statistical models (usually Gaussian i.i.d.) and a cost function \(C(x_k, \hat{x}_k)\).

Solutions are generally provided in the form:

\[
\hat{x}_k^o = G_k |Y_k|
\]

where \(G_k\) is the gain of the estimator.
<table>
<thead>
<tr>
<th></th>
<th>$C(x_k, \hat{x}_k)$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MMSE STSA</strong></td>
<td>$(x_k - \hat{x}_k)^2$</td>
</tr>
<tr>
<td>(Ephraim and Malah, 1984)</td>
<td></td>
</tr>
<tr>
<td><strong>MMSE Log-STSA (LSA)</strong></td>
<td>$(\log x_k - \log \hat{x}_k)^2$</td>
</tr>
<tr>
<td>(Ephraim and Malah, 1985)</td>
<td></td>
</tr>
<tr>
<td><strong>$\beta$-order MMSE STSA ($\beta$-SA)</strong></td>
<td>$(x_k^\beta - \hat{x}_k^\beta)^2$</td>
</tr>
<tr>
<td>(You et al., 2005)</td>
<td></td>
</tr>
<tr>
<td><strong>Weighted Euclidean (WE)</strong></td>
<td>$x_k^p (x_k - \hat{x}_k)^2$</td>
</tr>
<tr>
<td>(Loizou, 2005)</td>
<td></td>
</tr>
</tbody>
</table>
INTRODUCTION

BACKGROUND
Speech enhancement
Bayesian STSA estimation for speech enhancement

THESIS CONTRIBUTIONS
1 - Further analysis of the $\beta$-SA estimator
2 - Weighted $\beta$-SA with auditory-based parameter values
3 - Analytical generalization of Bayesian STSA estimators
4 - Multi-dimensional estimators allowing correlated frequency components

CONCLUSION
1 - **FURTHER ANALYSIS OF THE $\beta$-SA ESTIMATOR**

**Background**: $\beta$-SA $\Rightarrow$ $C(\chi_k, \hat{\chi}_k) = \left(\chi_k^\beta - \hat{\chi}_k^\beta\right)^2$

**Motivation**: The case $\beta < 0$ has not been studied in (You et al., 2005).

(i) - The $\beta$-SA cost function is normalized when $\beta < 0$:

$$C(\chi_k, \hat{\chi}_k, \beta) = \left(\chi_k^{-|\beta|} - \hat{\chi}_k^{-|\beta|}\right)^2 = \frac{C(\chi_k, \hat{\chi}_k; |\beta|)}{(\chi_k \hat{\chi}_k)^2|\beta|}$$

$\Rightarrow$ Favors an accurate estimation for small clean speech STSA, i.e. where the residual noise is less likely to be masked.
1 - FURTHER ANALYSIS OF THE $\beta$-SA ESTIMATOR

**Background:** $\beta$-SA $\Rightarrow$ $C(x_k, \hat{x}_k) = (x_k^\beta - \hat{x}_k^\beta)^2$

**Motivation:** The case $\beta < 0$ has not been studied in (You et al., 2005).

(i) - The $\beta$-SA cost function is normalized when $\beta < 0$:

$$C(x_k, \hat{x}_k, \beta) = \left(x_k^{-|\beta|} - \hat{x}_k^{-|\beta|}\right)^2 = \frac{C(x_k, \hat{x}_k; |\beta|)}{(x_k \hat{x}_k)^2|\beta|}$$

$\Rightarrow$ Favors an accurate estimation for small clean speech STSA, i.e. where the residual noise is less likely to be masked.
(ii) - $\beta < 0$ adds flexibility in terms of achievable $G_k$

(iii) - Proved analytically that the $\beta$-SA with $\beta \to 0$ tends to LSA estimator.
(ii) - $\beta < 0$ adds flexibility in terms of achievable $G_k$

(iii) - Proved analytically that the $\beta$-SA with $\beta \rightarrow 0$ tends to LSA estimator.
2 - $W_{\beta}$-SA WITH AUDITORY-BASED PARAMETER VALUES

**Motivation:** Combine the compression performed by the $\beta$-SA with the normalization of the WE and take advantage of the auditory interpretations that can be given to their parameters.

\[(1) - \text{W}_{\beta}$-SA ESTIMATOR\]

\[
\hat{x}_{k}^{W_{\beta}-SA} = \arg \min_{\hat{x}_{k}} E \left\{ \left( \frac{\chi_{k}^{\beta} - \hat{x}_{k}^{\beta}}{\chi_{k}^{\alpha}} \right)^{2} \right\}
\]

\[\Rightarrow \hat{x}_{k}^{W_{\beta}-SA} = G_{k}^{W_{\beta}-SA} |Y_k|\]

\[G_{k}^{W_{\beta}-SA} = \sqrt{\upsilon_k} \left( \frac{\Gamma \left( \frac{\beta - 2\alpha}{2} + 1 \right) M \left( -\frac{\beta - 2\alpha}{2}, 1; -\upsilon_k \right)}{\Gamma \left( -\alpha + 1 \right) M \left( \alpha, 1; -\upsilon_k \right)} \right)^{1/\beta}\]
2 - $W^\beta$-SA WITH AUDITORY-BASED PARAMETER VALUES

**Motivation:** Combine the compression performed by the $\beta$-SA with the normalization of the WE and take advantage of the auditory interpretations that can be given to their parameters.

(1) - $W^\beta$-SA ESTIMATOR

$$\hat{X}_k^{W^\beta-SA} = \arg \min_{\hat{X}_k} E \left\{ \left( \frac{X_\beta^k - \hat{X}_k^\beta}{X_\alpha^k} \right)^2 \right\}$$

$$\Rightarrow \hat{X}_k^{W^\beta-SA} = G_k^{W^\beta-SA} |Y_k|$$

$$G_k^{W^\beta-SA} = \frac{\sqrt{\nu_k}}{\gamma_k} \left( \frac{\Gamma \left( \frac{\beta-2\alpha}{2} + 1 \right) M \left( -\frac{\beta-2\alpha}{2}, 1; -\nu_k \right)}{\Gamma \left( -\alpha + 1 \right) M \left( \alpha, 1; -\nu_k \right)} \right)^{1/\beta}$$
CHOICE OF $\beta$ VALUES

(II) - CONSIDER TWO AUDITORY ASPECTS:

- Loudness $\beta = 1/3$
- Compressive nonlinearity $\beta = \beta_k$
CHOICE OF $\alpha$ VALUES

(III) - CONSIDER AUDITORY MASKING

- Smaller $X_k$ will not mask the residual noise.
- Since there is less speech energy at high frequencies:  
  $\Rightarrow$ Improve the estimation of small $X_k$ at high frequencies.
Proposed estimators found to be better than comparative estimators for:

- **Objective results**: $\text{SNR}_{\text{seg}}$, LLR, wPESQ
- **Subjective results**: MUSHRA

**Figure**: MUSHRA results (SNR = 0 dB).
3 - Analytical Generalization of Bayesian STSA Estimators

**Motivation:** Many existing Bayesian STSA estimators share a common structure.

(i) - GWSA Family of Estimators

\[
C_{GWSA}(\mathbf{x}_k, \hat{\mathbf{x}}_k) = \left( \frac{\mathbf{x}_k^\beta - \hat{\mathbf{x}}_k^\beta}{\mathbf{x}_k^\alpha \hat{\mathbf{x}}_k^\eta} \right)^2
\]

\[
\Rightarrow \hat{\mathbf{x}}_k = G_k|Y_k| \quad \text{where} \quad G_k = \sqrt{\nu_k} \left( \frac{-b' \pm \sqrt{b'^2 - 4a'c'}}{2a'} \right)^{\frac{1}{\beta}}
\]

where:

\[
a' = (\beta - \eta)\Gamma(-\alpha + 1)M(\alpha, 1; -\nu_k)
\]

\[
b' = (2\eta - \beta)\Gamma \left( \frac{\beta - 2\alpha}{2} + 1 \right) M \left( -\frac{\beta - 2\alpha}{2}, 1; -\nu_k \right)
\]

\[
c' = -\eta \Gamma(\beta - \alpha + 1) M(\alpha - \beta, 1; -\nu_k)
\]
3 - **ANALYTICAL GENERALIZATION OF BAYESIAN STSA ESTIMATORS**

**Motivation:** Many existing Bayesian STSA estimators share a common structure.

**(I) - GWSA FAMILY OF ESTIMATORS**

\[
C_{\text{GWSA}}(\hat{x}_k, \hat{x}_k) = \left( \frac{\hat{x}_k^\beta - \hat{x}_k^\beta}{\hat{x}_k^\alpha \hat{x}_k^\eta} \right)^2
\]

\[
\Rightarrow \hat{x}_k = G_k |Y_k| \quad \text{where} \quad G_k = \frac{\sqrt{\nu_k}}{\gamma_k} \left( \frac{-b' \pm \sqrt{b'^2 - 4a'c'}}{2a'} \right)^{\frac{1}{\beta}}
\]

where:  
\[
a' = (\beta - \eta)\Gamma(-\alpha + 1)M(\alpha, 1; -\nu_k)
\]
\[
b' = (2\eta - \beta)\Gamma(\frac{\beta-2\alpha}{2} + 1)M(-\frac{\beta-2\alpha}{2}, 1; -\nu_k)
\]
\[
c' = -\eta\Gamma(\beta - \alpha + 1)M(\alpha - \beta, 1; -\nu_k)
\]
(ii) - GWSA generalizes many existing Bayesian STSA estimators:

<table>
<thead>
<tr>
<th></th>
<th>( C(\hat{x}_k, \hat{x}_k) )</th>
<th>( \beta )</th>
<th>( \alpha )</th>
<th>( \eta )</th>
</tr>
</thead>
<tbody>
<tr>
<td>MMSE STSA (Ephraim and Malah, 1984)</td>
<td>((\hat{x}_k - \hat{x}_k)^2)</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>LSA (Ephraim and Malah, 1985)</td>
<td>((\log x_k - \log \hat{x}_k)^2)</td>
<td>(\to 0)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>COSH (Loizou, 2005)</td>
<td>(\frac{1}{2} \left( \frac{x_k}{\hat{x}_k} + \frac{\hat{x}_k}{x_k} \right) - 1)</td>
<td>1</td>
<td>0.5</td>
<td>0.5</td>
</tr>
<tr>
<td>WE (Loizou, 2005)</td>
<td>(x_k^p (x_k - \hat{x}_k)^2)</td>
<td>1</td>
<td>(-p/2)</td>
<td>0</td>
</tr>
<tr>
<td>WCOSH (Loizou, 2005)</td>
<td>(\left( \frac{x_k}{\hat{x}_k} + \frac{\hat{x}_k}{x_k} - 1 \right) x_k^q)</td>
<td>1</td>
<td>(\frac{(1-q)}{2})</td>
<td>0.5</td>
</tr>
<tr>
<td>(\beta)-SA (You, 2005)</td>
<td>((x_k^\beta - \hat{x}_k^\beta)^2)</td>
<td>(\beta)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(W\beta)-SA</td>
<td>(\left( \frac{x_k^\beta}{\hat{x}_k^\alpha} \right)^2)</td>
<td>(\beta)</td>
<td>(\alpha)</td>
<td>0</td>
</tr>
</tbody>
</table>
(iii) - \( \eta \) adds flexibility in terms of achievable gain curves when compared to existing estimators:

(iv) - All the estimators belonging to that family tend to a Wiener filter at high instantaneous SNRs.
(iii) - \( \eta \) adds flexibility in terms of achievable gain curves when compared to existing estimators:

(iv) - All the estimators belonging to that family tend to a Wiener filter at high instantaneous SNRs.
4 - Multi-dimensional Estimators Allowing Correlated Frequency Components

**Motivation:** correlation does exist between the different frequency components.
Adopted model: \( \mathbf{Y} = \mathbf{X} + \mathbf{W} \)

**PROPOSED ESTIMATOR**

\[
\hat{\mathbf{X}}^o = \arg\min_{\hat{\mathbf{X}}} E\{\| \mathbf{X} - \hat{\mathbf{X}} \|^2 \} = E\{\mathbf{X}|\mathbf{Y}\}
\]

- Considers all frequencies simultaneously.
- A closed-form expression for \( E\{\mathbf{X}|\mathbf{Y}\} \) is not readily available.
  - \( \Rightarrow \) We approach the problem by finding tractable upper and lower bounds on \( E\{\mathbf{X}|\mathbf{Y}\} \).
PROPOSED MULTIDIMENSIONAL ESTIMATORS

\[ \hat{X}^o = \arg\min_{\hat{X}} E\{\|X - \hat{X}\|^2\} = E\{X|Y\} \]

(i) - Lower bound on \( E\{X|Y\} \)

\[ \hat{X}_L^o = |G_{\text{MMSE}}Y| \quad \text{where} \quad G_{\text{MMSE}} \triangleq R_x(R_x + R_w)^{-1} \]

(ii) - Upper bound on \( E\{X|Y\} \)

\[ \hat{X}_U^o = (|G_{\text{MMSE}}Y|^2 + \text{diag}\{G_{\text{MMSE}}R_w\})^{1/2} \]

(iii) - Family of estimators

\[ \hat{X}_\gamma^o = (|G_{\text{MMSE}}Y|^2 + \gamma \text{diag}\{G_{\text{MMSE}}R_w\})^{1/2} \]
Proposed estimators found to be better than compared estimators for informal listening experiments, LLR and wPESQ.

**TABLE:** wPESQ results for white, pink and cockpit noises at several SNRs (10, 15 and 20 dB).

<table>
<thead>
<tr>
<th></th>
<th>MMSE</th>
<th>Wiener</th>
<th>MMSE</th>
<th>$\hat{X}_\gamma$</th>
<th>$\hat{X}_{\delta \text{MMSE}}$</th>
<th>$\hat{X}_\gamma$</th>
<th>$\hat{X}_{\delta \gamma}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>STSA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 dB</td>
<td>1.35</td>
<td>1.53</td>
<td>1.57</td>
<td>1.52</td>
<td>1.61</td>
<td>1.59</td>
<td></td>
</tr>
<tr>
<td>15 dB</td>
<td>1.70</td>
<td>1.90</td>
<td>1.94</td>
<td>1.98</td>
<td>1.98</td>
<td>2.11</td>
<td></td>
</tr>
<tr>
<td>20 dB</td>
<td>2.25</td>
<td>2.45</td>
<td>2.39</td>
<td>2.52</td>
<td>2.48</td>
<td>2.65</td>
<td></td>
</tr>
<tr>
<td>Pink</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 dB</td>
<td>1.47</td>
<td>1.58</td>
<td>1.70</td>
<td>1.74</td>
<td>1.71</td>
<td>1.77</td>
<td></td>
</tr>
<tr>
<td>15 dB</td>
<td>1.90</td>
<td>1.95</td>
<td>2.05</td>
<td>2.20</td>
<td>2.06</td>
<td>2.23</td>
<td></td>
</tr>
<tr>
<td>20 dB</td>
<td>2.48</td>
<td>2.48</td>
<td>2.49</td>
<td>2.66</td>
<td>2.55</td>
<td>2.72</td>
<td></td>
</tr>
<tr>
<td>Cockpit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10 dB</td>
<td>1.35</td>
<td>1.38</td>
<td>1.47</td>
<td>1.53</td>
<td>1.47</td>
<td>1.53</td>
<td></td>
</tr>
<tr>
<td>15 dB</td>
<td>1.69</td>
<td>1.65</td>
<td>1.77</td>
<td>1.92</td>
<td>1.76</td>
<td>1.91</td>
<td></td>
</tr>
<tr>
<td>20 dB</td>
<td>2.20</td>
<td>2.11</td>
<td>2.19</td>
<td>2.38</td>
<td>2.22</td>
<td>2.39</td>
<td></td>
</tr>
</tbody>
</table>
### INTRODUCTION

### BACKGROUND
Speech enhancement
Bayesian STSA estimation for speech enhancement

### THESIS CONTRIBUTIONS
1 - Further analysis of the $\beta$-SA estimator
2 - Weighted $\beta$-SA with auditory-based parameter values
3 - Analytical generalization of Bayesian STSA estimators
4 - Multi-dimensional estimators allowing correlated frequency components

### CONCLUSION
SUMMARY OF CONTRIBUTIONS

1 - Studied $\beta < 0$ in the $\beta$-SA estimator (Chapter 4):
   $\Rightarrow \beta$-SA cost function is normalized when $\beta < 0$
   $\Rightarrow \beta < 0$ adds flexibility in terms of achievable $G_k$.
   $\Rightarrow \beta$-SA with $\beta \to 0$ tends to LSA estimator

2 - Proposed $W\beta$-SA estimator (Chapter 5):
   $\Rightarrow$ Developed the $W\beta$-SA estimator combining power law and weighting factor.
   $\Rightarrow$ Choosing the parameters based on auditory considerations suggests decrease of $G_k$ at high frequencies which improves the noise reduction while limiting the speech distortion.
   $\Rightarrow$ Demonstrates a noticeable advantage over existing estimators, especially at low SNRs.
SUMMARY OF CONTRIBUTIONS

3 - Proposed GWSA estimator (Chapter 6):
⇒ Has many existing estimators as particular cases.
⇒ $\eta$ has an added flexibility in terms of achievable gain curves.
⇒ All the estimators belonging to that family tend to a Wiener filter.

4 - Proposed family of Bayesian estimators that assumes correlated frequency components (Chapter 7):
⇒ Developed closed-form solution for lower and upper bounds on the desired estimator.
⇒ Proposed a family of multidimensional estimators.
⇒ Demonstrates noticeable advantage over existing estimators at high SNRs.
RELATD PUBLICATIONS

Journals


**RELATED PUBLICATIONS**

**Conferences**

- **E. Plourde and B. Champagne.**
  Bayesian spectral amplitude estimation for speech enhancement with correlated frequencies.

- **E. Plourde and B. Champagne.**
  Perceptually based speech enhancement using the weighted $\beta$-SA estimator.
  *ICASSP’08*, Las Vegas, NV, Apr. 2008

- **E. Plourde and B. Champagne.**
  Integrating the cochlea’s compressive nonlinearity in the Bayesian approach for speech enhancement.
  *EUSIPCO’07*, Poznan, Poland, Sep. 2007

- **E. Plourde and B. Champagne.**
  Further Analysis of the $\beta$-Order MMSE STSA Estimator for Speech Enhancement.
  *CCECE’07*, Vancouver, BC, April 2007

- **E. Plourde and B. Champagne.**
  A family of Bayesian STSA estimators for the enhancement of speech with correlated frequency components.
  *ICASSP’10*, Dallas, TX, USA, submitted.
THANK YOU!