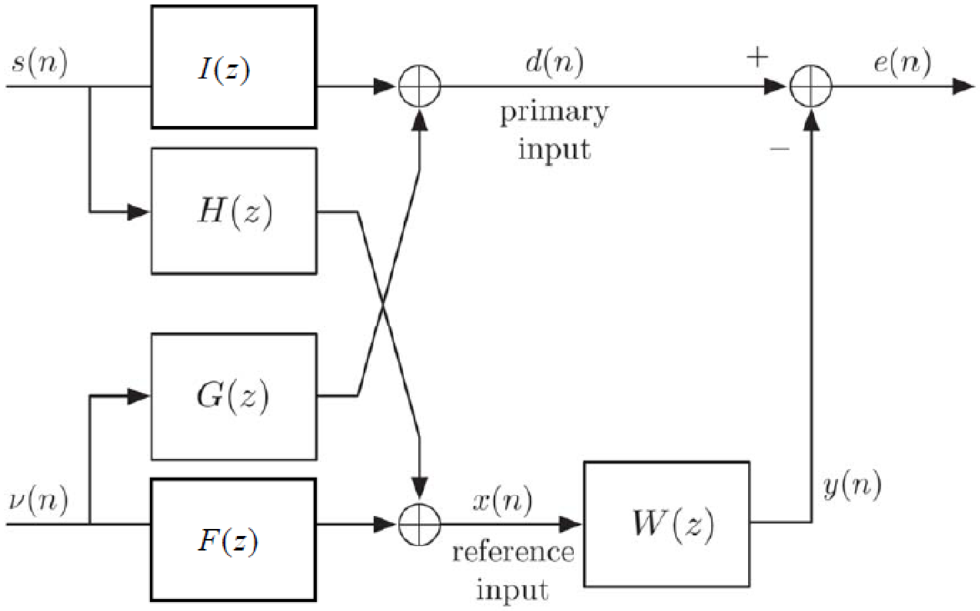
Consider the following noise canceler setup, where the objective is to remove from $e(n)$ the signal component correlated to $v(n)$, and only keep the component coming from $s(n)$:



with:

$$F(z)=\frac{1}{1-0.95z^{-1}} \quad |z|>0.95 \qquad G(z)=5+\frac{5}{3}z^{-1} \qquad H(z)=0 \qquad I(z)=3+\frac{3}{4}z^{-1}\ .$$

Using a $N = 3$ taps adaptive FIR filter $\boldsymbol{w}(k)$ with initial conditions $\boldsymbol{w}(0)=\mathbf{0}$, plot the normalized MSE function $\frac{\xi(\boldsymbol{w}(k))}{\sigma_d^2}=\frac{\sigma_d^2-2\boldsymbol{w}^T(k)\boldsymbol{p}+\boldsymbol{w}^T(k)\boldsymbol{R}\boldsymbol{w}(k)}{\sigma_d^2}$ over the iteration number k, i.e., the learning curve, for the steepest descent and Newton algorithms:

$\boldsymbol{w}(k+1)=\boldsymbol{w}(k)-\mu(\boldsymbol{R}\,\boldsymbol{w}(k)-\boldsymbol{p})$ (steepest descent)

$\boldsymbol{w}(k+1)=\boldsymbol{w}(k)-\mu\boldsymbol{R}^{-1}(\boldsymbol{R}\,\boldsymbol{w}(k)-\boldsymbol{p})$ (Newton),

and consider the following two cases:

- The theoretical values of $\boldsymbol{p}$ and $\boldsymbol{R}$ are derived from the knowledge of the transfer functions (above) and assuming that $s(n)$ and $v(n)$ are sample sequences from white noise zero-mean unit-variance random processes.

- $\boldsymbol{p}$ and $\boldsymbol{R}$ are estimated (measured) using all the samples of $x(n)$ and $d(n)$ provided in the .wav sound files, where $d(n)$ is a mixture of components from the male voice $s(n)$ and female voice $v(n)$, while $x(n)$ only includes a component from the female voice $v(n)$. For this case, you also need to generate the $e(n)$ signal obtained by filtering $x(n)$ with the final $\boldsymbol{w}$ coefficients obtained from the Newton method, and save $e(n)$ as a .wav file (to submit with your report).

Use 500 iterations, with $\mu=0.25/(N\sigma_x^2)$ for the steepest descent method and $\mu=0.25$ for the Newton method. Note that σ_x^2 will be different between the two considered cases above, so μ for the steepest descent will also be different.

Moreover, the convergence speed of the steepest descent will be different between the two cases, because the eigenvalue spread in the input signal $x(n)$ of the adaptive filter will be larger in the 2nd case (female speech source for $v(n)$ instead of white noise source, so larger eigenvalue spread). The MSE function $\xi(\boldsymbol{w}(k))=\sigma_d^2-2\boldsymbol{w}^T(k)\boldsymbol{p}+\boldsymbol{w}^T(k)\boldsymbol{R}\boldsymbol{w}(k)$ is also a bit different between the two cases (different $\sigma_d^2, \boldsymbol{p}, \boldsymbol{R}$ values, because different $v(n)$ and $s(n)$ sources).