Proof of Theorem

I. THE SOLUTION FOR (7)

The constrained optimization problem of the MGB is:

$$\mathbf{g}^*(t) = \min_{\mathbf{g}(t)} \frac{1}{2}||\hat{\mathbf{g}}_k(t) - \mathbf{g}(t)||_2^2 \tag{1}$$
$$s.t. \quad \tilde{\mathbf{G}}(t)^\mathrm{T}\mathbf{g}(t) \geq 0,$$

where k is the ordinal number of the complex language, $\hat{\mathbf{g}}_k(t) = w_k(t+1)\mathbf{g}_k(t)$ is the normalized gradient, $\mathbf{g}^*(t)$ is the balancing gradient, and $\tilde{\mathbf{G}}(t)$ is the set of the rest normalized gradients. We can obtain the lagrangian function of (1), which is shown as follows:

$$L(\mathbf{g}(t), \boldsymbol{\gamma}) = \frac{1}{2}\mathbf{g}(t)^\mathrm{T}\mathbf{g}(t) - \hat{\mathbf{g}}_k(t)^\mathrm{T}\mathbf{g}(t) - \boldsymbol{\gamma}\tilde{\mathbf{G}}(t)^\mathrm{T}\mathbf{g}(t), \tag{2}$$

where $L(\mathbf{g}(t), \boldsymbol{\gamma})$ is the lagrangian function and $\boldsymbol{\gamma}$ is the lagrange multiplier.

The lagrangian dual form of (1) is shown as follows:

$$f(\boldsymbol{\gamma}) = \min_{\mathbf{g}(t)} L(\mathbf{g}(t), \boldsymbol{\gamma}) \tag{3}$$
$$s.t. \quad \boldsymbol{\gamma} \geq 0,$$

where $f(\boldsymbol{\gamma})$ is the lagrangian dual function.

We next find $\mathbf{g}^*(t)$ that minimizes (2) by setting the derivative of (2) to zero:

$$\frac{L(\partial\mathbf{g}(t), \boldsymbol{\gamma})}{\partial\mathbf{g}(t)} = \mathbf{g}(t)^\mathrm{T} - \hat{\mathbf{g}}_k(t)^\mathrm{T} - \boldsymbol{\gamma}\tilde{\mathbf{G}}(t)^\mathrm{T} = 0,$$
$$\implies \mathbf{g}^*(t) = \hat{\mathbf{g}}_k(t) + \tilde{\mathbf{G}}(t)\boldsymbol{\gamma}^\mathrm{T}. \tag{4}$$

We further put $\mathbf{g}^*(t)$ in (3), and it can be written as:

$$f(\boldsymbol{\gamma}) = \frac{1}{2}(\hat{\mathbf{g}}_k(t)^\mathrm{T}\hat{\mathbf{g}}_k(t) + 2\boldsymbol{\gamma}\tilde{\mathbf{G}}(t)^\mathrm{T}\hat{\mathbf{g}}_k(t) + \boldsymbol{\gamma}\tilde{\mathbf{G}}(t)^\mathrm{T}\tilde{\mathbf{G}}(t)\boldsymbol{\gamma}^\mathrm{T})$$
$$- \hat{\mathbf{g}}_k(t)^\mathrm{T}\hat{\mathbf{g}}_k(t) - 2\boldsymbol{\gamma}\tilde{\mathbf{G}}(t)^\mathrm{T}\hat{\mathbf{g}}_k(t) - \boldsymbol{\gamma}\tilde{\mathbf{G}}(t)^\mathrm{T}\tilde{\mathbf{G}}(t)\boldsymbol{\gamma}^\mathrm{T}) \tag{5}$$
$$= -\frac{1}{2}\hat{\mathbf{g}}_k(t)^\mathrm{T}\hat{\mathbf{g}}_k(t) - \boldsymbol{\gamma}\tilde{\mathbf{G}}(t)^\mathrm{T}\hat{\mathbf{g}}_k(t) - \frac{1}{2}\boldsymbol{\gamma}\tilde{\mathbf{G}}(t)^\mathrm{T}\tilde{\mathbf{G}}(t)\boldsymbol{\gamma}^\mathrm{T}$$

The solution $\boldsymbol{\gamma}^* = \max_{\boldsymbol{\gamma};\boldsymbol{\gamma}>0} f(\boldsymbol{\gamma})$ to the dual form is obtained by:

$$\frac{\partial f(\boldsymbol{\gamma})}{\partial\boldsymbol{\gamma}} = -\hat{\mathbf{g}}_k(t)^\mathrm{T}\tilde{\mathbf{G}}(t) - \boldsymbol{\gamma}\tilde{\mathbf{G}}(t)^\mathrm{T}\tilde{\mathbf{G}}(t) = 0,$$
$$\implies \boldsymbol{\gamma}^* = -\frac{\hat{\mathbf{g}}_k(t)^\mathrm{T}\tilde{\mathbf{G}}(t)}{\tilde{\mathbf{G}}(t)^\mathrm{T}\tilde{\mathbf{G}}(t)}. \tag{6}$$

Finally, because $\boldsymbol{\gamma}^*$ must be grater than zeros, we obtain the balancing gradient by putting $\boldsymbol{\gamma}^*$ in (4):

$$\mathbf{g}^*(t) = \hat{\mathbf{g}}_k(t) - \tilde{\mathbf{G}}(t)[\frac{\hat{\mathbf{g}}_k(t)^\mathrm{T}\tilde{\mathbf{G}}(t)}{\tilde{\mathbf{G}}(t)^\mathrm{T}\tilde{\mathbf{G}}(t)}]_-^\mathrm{T}. \tag{7}$$