

3.1 Dans les figures 4 et 5, les droites sont neurones, le nombre de neurones est égale au nombre de droites, et la forme de la surface de réseau neurones se compose de différents droites et ces droites sont combinaisons linéaires des neurones.

La discriminateur RN améliore sa surface à partir de discriminateur quadratique.

Quand on augmente  $P_{gen}$ , le  $\tau$  de 3 discriminateurs ne change pas presque. Quand on change  $\eta$  de 0.01 au 0.1 ou de 0.01 au 0.001, dans un cas le discriminateur RN ne peut pas fonctionner bien, et le  $\tau_{app}$  et  $\tau_g$  ne sont pas améliorés, et  $\eta$  correspond au largement de variation de  $J$  et  $\tau$ .

Quand on augmente  $N_c$ , les  $\tau$  et  $J$  ne changent pas presque à partir de quelque  $N_c$  (pas très grand), seulement la frontière est plus lisse.

Quand  $P_{app}$  augmente le  $\tau_{app}$  et  $\tau_g$  sont convergés au  $\tau_g$  ( $\mu$  et  $\gamma$  connue).

3.2 les performances de discriminateur ( $\tau_g$ ) ne change pas selon le nombre de chiffres, mais quand  $P_{gen}$  est fixé la covariance de  $\tau_g$  est constante,  $P_{gen}$  augmente, covariance diminue,  $\tau_g$  est plus précis.

3.3 Pour discriminateur RN, quand le nombre de chiffres augmente,  $\tau_g$  de RN diminue. Car il faut distinguer plus de chiffres.

3.4 Quand on génère les bases de généralisation d'après l'équation 49, les deux filtres adaptés n'ont pas une grande différence sur  $\tau_g$ , sont tous près 0.9, par contre, quand on génère les bases selon 51, le  $\tau_g$  de filtre FA gaussien (en sachant l'amplitude) est beaucoup plus petite que celui d'amplitude inconnu. Car quand les bases sont en même amplitude, les deux filtres sont tous adaptés, mais quand le cas devient plus complexe, les bases ne sont pas en même amplitude, le premier filtre n'est plus adapté. Ici, on trouve que "qui peut le plus peut le moins" est vrai. Le deuxième filtre est plus capable.

3.5 Dans le premier cas,  $\tau_{app}$  et  $\tau_g$  tendent vers ses limitations rapidement (avec peu de nombre de passes),  $\tau_{app}$  tend vers 1, par contre  $\tau_g$  tend vers presque 0.5 (beaucoup plus bas que  $\tau_{app}$ ). C'est parce que la base d'apprentissage est formée par le modèle l'équation (49) qui manque un paramètre  $A$  (Amplitude) que l'équation (51), ainsi le RN qui est entraîné par une base qui manque un paramètre est bien sûr non-adapté à une base de généralisation plus complexe.

Dans le second cas,  $\tau_{app}$  et  $\tau_g$  tendent vers ses limitations moins rapidement que dans le premier cas, car la base est plus complexe. Et cette fois-ci, la limitation de  $\tau_g$  augmente fortement, car la base de généralisation et la base d'apprentissage sont formées par le même modèle.

Quand le nombre de chiffres augmente, la diversité d'une base augmente, donc il prend plus de nombre de passes pour atteindre la limitation du  $\tau_g/\tau_{app}$ . En même temps, les limitations de  $\tau_{app}$  et  $\tau_g$  diminuent, car la reconnaissance est plus complexe, il faut distinguer plus de nombres, donc la validation diminue.

Pour le cas moins complexe, par exemple le nombre de chiffres est petit, ou bien le modèle est simple (on ne considère pas d'amplitude), la performance du RN est mieux que le discriminateur d'amplitude inconnu. Mais pour le cas complexe, le discriminateur d'amplitude est mieux, or on doit ajouter plus de neurones.