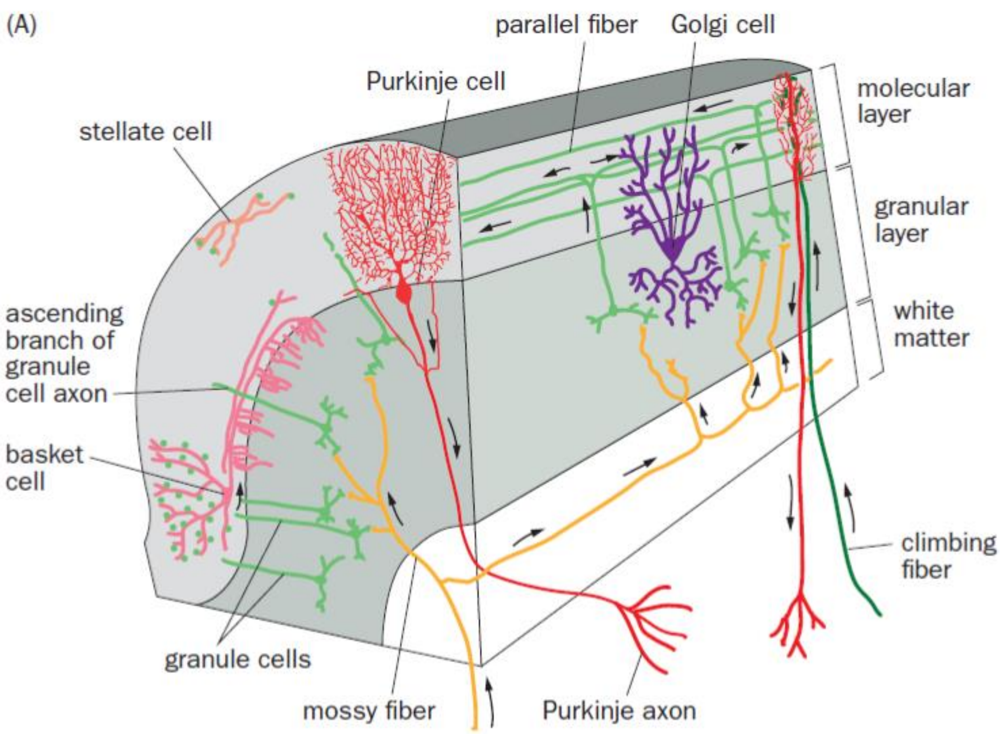
How granule cells sample inputs

Due Wednesday, Sep 26, 2018

Du Yihui,　　PB16001406

算法说明：



N: the number of granule cells

M:the number of mossy fiber inputs

K: the convergence of a granule cell

对于给定的 N 和 M, $\text{combination(K)} = \binom{M}{K}$. 所有的 granule cells 得到不同输入组合的概率为 p(K)=$\frac{A^N_{\text{combination(K)}}}{\text{combination(K)}^N}$，其中 $A^N_{\text{combination(K)}}$ 为排序数(即：将N 个 granule cells 看做不同细胞，当然也可看成相同细胞，则 p(K)=$\frac{C^N_{\text{combination(K)}}}{\text{combination(K)}^N/N!}$)，结果一致)。

Ⅰ. What is the probability p that granule cells all receive different combinations of inputs?

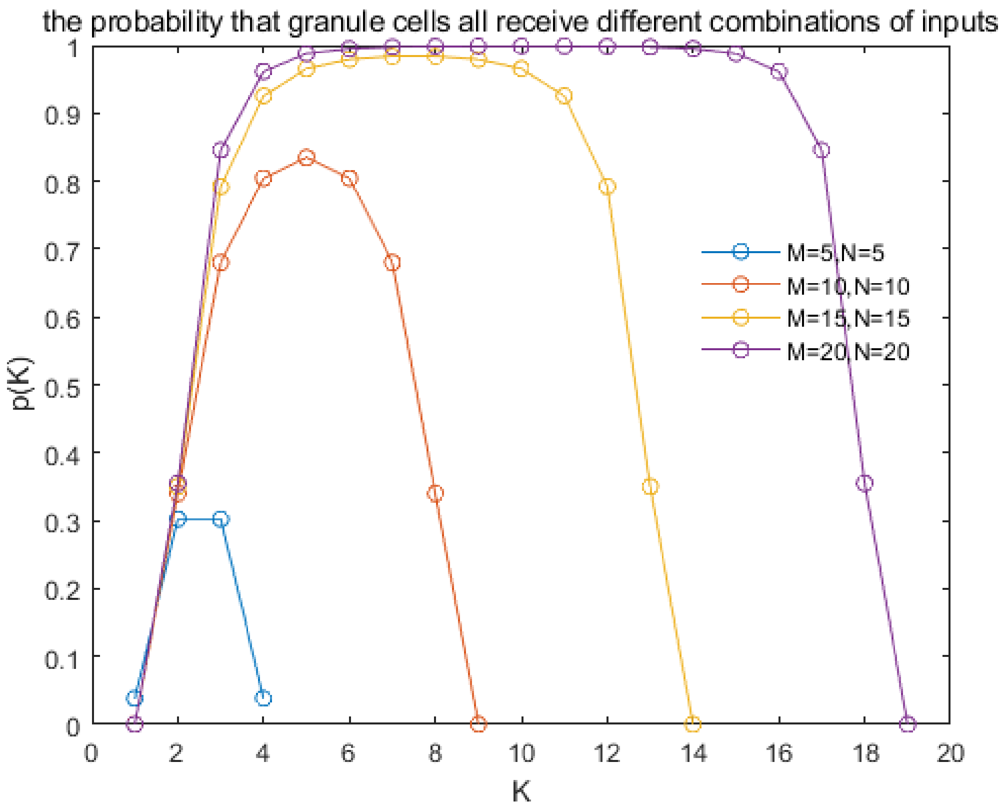
$$p(K)=\frac{A^{N}_{\text{combination(K)}}}{\text{combination(K)}^{N}}, \quad \text{combination(K)} = \binom{M}{K}.$$

Ⅱ. For given N and M, plot p as a function of K, and show when p reaches its maximum.

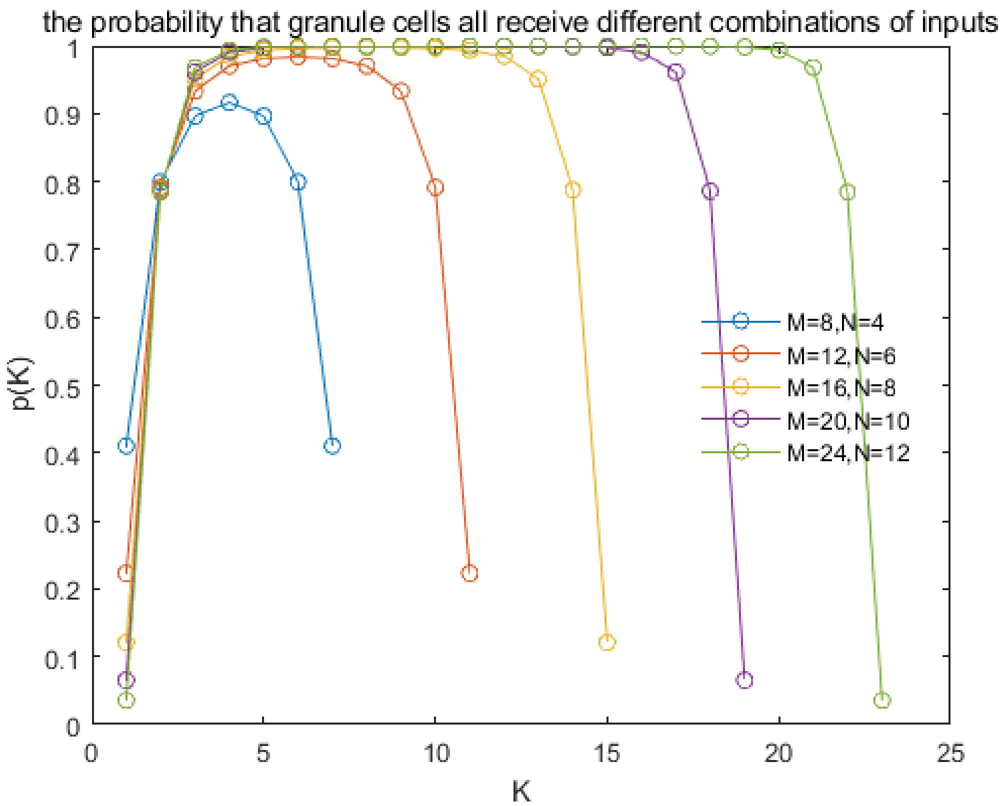
Code:

```
figure;
for N= 2:2:10
    for M = N:N:3*N;
        for K=1:1:M-1
            Combinations(K)=nchoosek(M,K);
            p(K)=prod((Combinations(K)-N+1):Combinations(K))/(Combinations(K)^N);
        end
        plot(1:1:M-1,p,'o-');
        clear p;
        hold on
    end
end
title('the probability that granule cells all receive different combinations of inputs');
xlabel('K');
ylabel('p(K)');
legend('M=2,N=2','M=4,N=2','M=6,N=2','M=4,N=4','M=8,N=4','M=12,N=4','M=6,N=6','M=12,N=6','M=18,N=6'
legend('Location','best');
legend('boxoff');
```

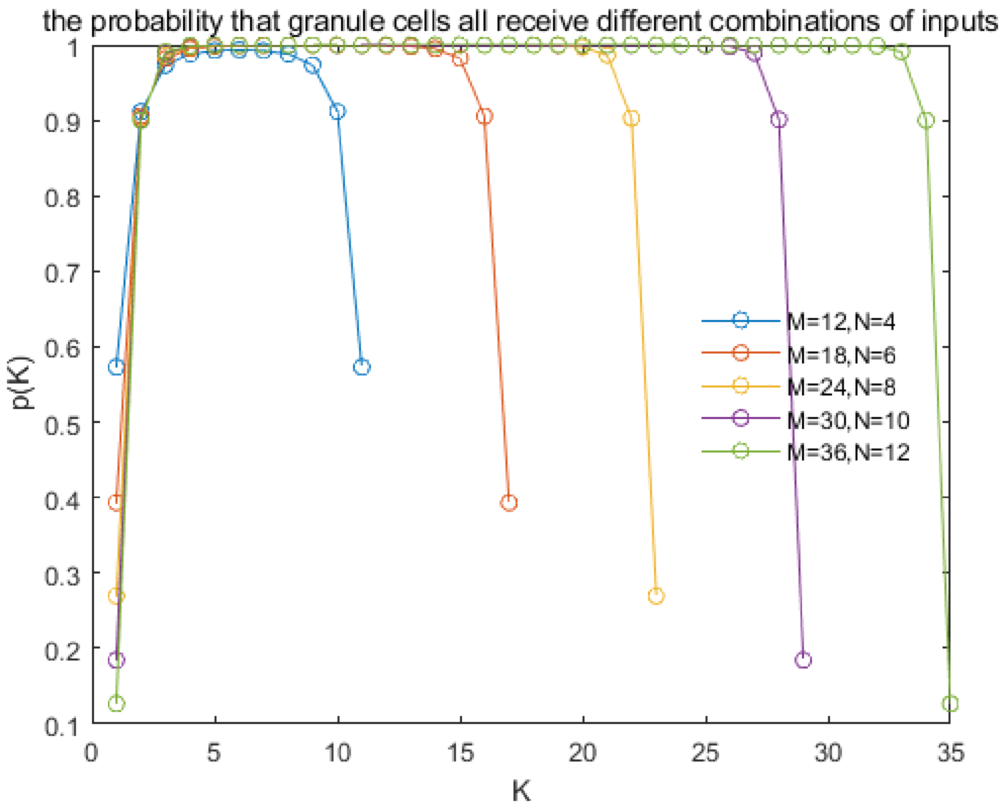
Result:



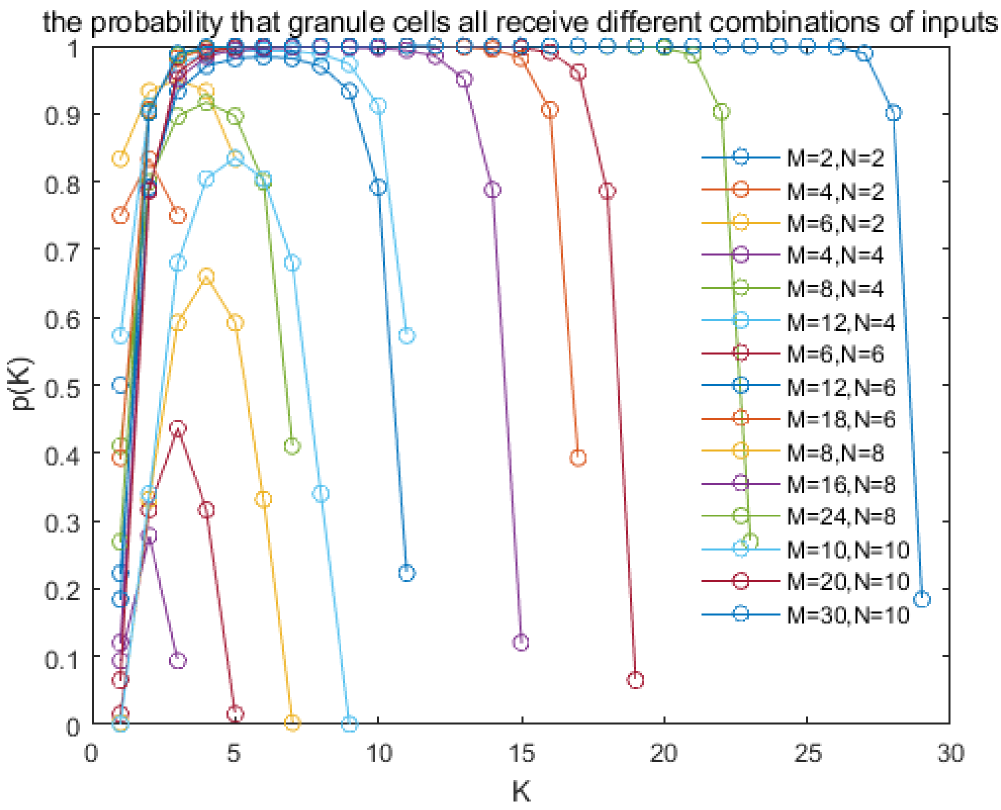
the probability that granule cells all receive different combinations of inputs
p(K)
K
M=5,N=5
M=10,N=10
M=15,N=15
M=20,N=20



the probability that granule cells all receive different combinations of inputs
p(K)
K
M=8,N=4
M=12,N=6
M=16,N=8
M=20,N=10
M=24,N=12



the probability that granule cells all receive different combinations of inputs
p(K)
K
M=12,N=4
M=18,N=6
M=24,N=8
M=30,N=10
M=36,N=12



the probability that granule cells all receive different combinations of inputs
p(K)
K
M=2,N=2
M=4,N=2
M=6,N=2
M=4,N=4
M=8,N=4
M=12,N=4
M=6,N=6
M=12,N=6
M=18,N=6
M=8,N=8
M=16,N=8
M=24,N=8
M=10,N=10
M=20,N=10
M=30,N=10

Analysis:

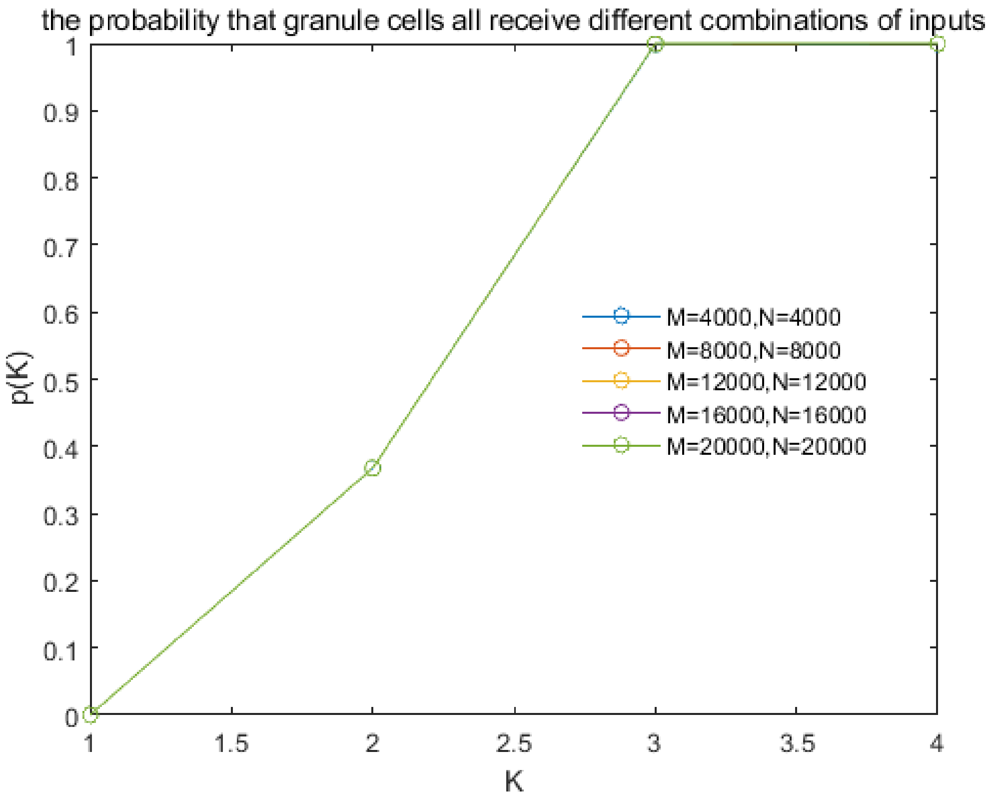
对于较小的 N 和 M，combination(K) $= \binom{M}{K}$.的量级较小，可以进行数值运算，结果如上。可以看出，对于不同的 N,M，p 均在 K 较小时收敛于 1。因此，对于较大的 N 和 M，只能进行符号运算，只需运算前几位 K 即可。不过就结果来看，matlab 的符号运算结果也有些问题。

Code:

```
figure;
for N= 4000:4000:20000
        M = N;
        for K=1:1:4
            Combinations(K)=nchoosek(sym(M),sym(K));
            p(K)=prod(sym((Combinations(K)-N+1):Combinations(K)))/(sym(Combinations(K))^N);
        end
        plot(1:1:4,p,'o-');
        %clear p;
        hold on

end
title('the probability that granule cells all receive different combinations of inputs');
xlabel('K');
ylabel('p(K)');
legend('M=4000,N=4000','M=8000,N=8000','M=12000,N=12000','M=16000,N=16000','M=20000,N=20000');
legend('Location','best');
legend('boxoff');
```

Result:

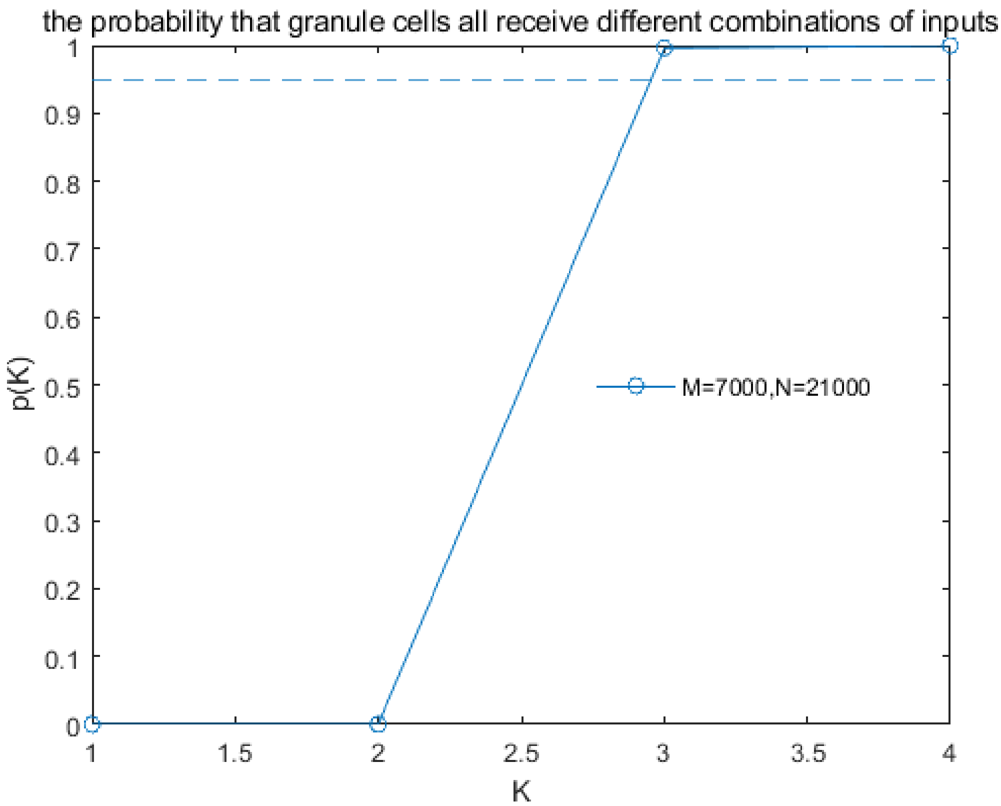


Ⅲ. Using N = 21000, M = 7000, compute K when p approaches 95 percent of its maximum.

Code:

```
N = 21000;
M = 7000;
figure;
for K=1:1:4
    Combinations(K)=nchoosek(sym(M),sym(K));
    p(K)=prod(sym((Combinations(K)-N+1):Combinations(K)))/(sym(Combinations(K))^N);
end
plot(1:1:4,p,'o-');
hold on
line([1,4],[0.95,0.95],'linestyle','--');
title('the probability that granule cells all receive different combinations of inputs');
xlabel('K');
ylabel('p(K)');
legend('M=7000,N=21000');
legend('Location','best');
legend('boxoff');
```

Result:



可以看出，K=3 时，p 已大于 0.95. 但是感觉符号计算的误差较大。

Ⅳ. Discuss whether it is beneficial to have small K when M is very large.

我认为答案是肯定的。

当 M 很大时，K 很小时，p 就可以接近 1。也就是说，只需要很小的 K，就可以保证所有的 granule cells 得到不同输入组合，即得到不同的信息输入。且 K 越小，细胞处理信息的效率越高。