(I)

1. The notation $p_e(r|x_j)$ is not good.

2. A simpler reasoning:

(a) For any realization W there are equiprobable realizations that correspond to permuting stimuli $\Longrightarrow$

$$\underset{\text{averaged over } W}{MSE} = P_{error} \times \underbrace{\frac{1}{L}\sum_{i=1}^{L}\left\langle (x_i - \hat{x}_i)^2 \,\middle|\, \hat{x}_i \neq x_i \right\rangle_W}_{\overline{e^2}} .$$

<u>and</u> $P(\hat{x}_i = x_j \mid \hat{x}_i \neq x_i) = \frac{1}{L-1} \quad \forall j \neq i .$

(b) $P_{error} = 1 - P_{no\ error}$

$= 1 - (\dots)^{L-1}$ etc.

(c) For $\overline{e^2}$:

* $\left\langle (x_i - \hat{x}_i)^2 \right\rangle = \frac{1}{L-1} \sum_{\substack{j=1 \\ j\neq i}}^{L} \left(\frac{i-j}{L}\right)^2$

* $\overline{e^2} = \frac{1}{L\cdot(L-1)} \sum_{i \neq j} \left(\frac{i-j}{L}\right)^2$

3. Don't use "in the first step...". Either "the equality is obtained by..." or something like this; or reformulate differently.

(II)

1. The sections "Local error" and "Global error" should be <u>sub</u>-sections of "Broad timing curves" (not sections on the same level).

2. Instead of defining local & global errors, say that here we calculate the MSE in the case in which the response curve is smooth on the scale of the noise (large enough σ, write the inequality). Then errors can be with nearby stimulus (local) or with distant stimulus because resp. curve loops back (global).

$$MSE = MSE_{local} + MSE_{global}$$

3. $MSE_{local} = \int_0^1 dx \, \frac{\eta^2}{[v'(x)]^2}$ and

$$[v'(x)]^2 = \sum_{i,j,j'} W_{ij} W_{ij'} \, u'(x-c_j) u'(x-c_{j'})$$

$$\approx \sum_{i,j} (W_{ij})^2 [u'(x-c_j)]^2 \quad \text{(for large } L\text{)}$$

Now, by central-limit theorem: $\sum_j (W_{ij})^2 [u'(x-c_j)]^2$ is a Gaussian rand. var. with mean $\approx$

$$\sum_i L \cdot \frac{1}{L} \cdot \int_0^1 dy \, \frac{(x-y)^2}{\sigma^4} A^2 \exp\left(-\frac{(x-y)^2}{\sigma^2}\right) = \frac{\sqrt{\pi}}{2} \frac{NA^2}{\sigma}$$

& calculate variance; $\boxed{\text{but check \& mention the Lindeberg condition}}$.

4. For MSE_{global}, I would simply say that we approximate the response curve by a set of $\frac{1}{6}$ uncorrelated segments. Then the only thing that changes in the previous derivation is the prefactor : $L \to \frac{1}{6}$.

III

1. Call this section "Influence of input noise on population coding" and place it __after__ the section on multi-dimensional stimuli.

2. This identity is always true. L large implies that the second term is small; reformulate.

3. It is __not__ a development in $\frac{N}{L}$; furthermore, it is not clear whether you do a development in $\frac{1}{L}$ or in $\frac{\xi}{\tilde{\eta}}$. The text suggest that it is the former (and indeed there is no reason that it be the latter since ξ & η can be of the same order). __But__ this seems contradicted by your results, since both 1st & 2nd order terms are of order $\frac{N^2}{L^2}$. So is it really a correct expansion? Are all order terms scaling like $\frac{N^2}{L^2}$?

4. Problem related to point 3 above: in a correct expansion, with $J \approx 1+\varepsilon+\varepsilon^2$, for $\frac{1}{J}$ you would need to include __both__ 1st-order correction in $\varepsilon+\varepsilon^2$ __and__ second-order correction in ε.

IV

1. All the definitions/model explanations should come in the earlier section.
2. Define clearly the MSE. For a multi-dimensional stimulus, there can be various definitions.