1 Mike Questions

1. Is there a difference between $\vec{c_i}$ and c_i? Should we just use $\vec{c}$ and c_i?
2. Why does a $\vec{c_i}$ term appear in the fitness? is ω_{c_i} a function of $\vec{c_i}$ in this notation?
3. $\vec{c}$ denotes the current gene, correct? Or does it represent the entire genome?
4. If $\vec{c}$ does denote the whole genome, should ϕ have some sort of subscript relating it to the current gene in the genome?
5. Is it alright if I use ν to denote the number of synonyms to $\vec{c_i}$?
6. Where do the mutation terms go? I think I'll include them for completeness. Just turn $\exp\left[-\sum_j \beta_{c_j}\omega_{c_j}(\vec{c_j})\phi\right]$ into $\exp\left[-\sum_j(\mu_{c_j} - \beta_{c_j}\omega_{c_j}(\vec{c_j}))\phi\right]$?
7. $\mathbb{C}$ only contains synonymous sequences, correct?
8. The last cancellation strikes me as odd, but I'm working through it now.

2 Notation

$\vec{c}$ is the vector of codons (of length n, the length of the ??gene/genome??)
??$\vec{c_i}/c_i$?? is the i^{th} codon of $\vec{c}$
$\vec{c}_{i\ell}$ is the ℓ^{th} synonym of codon c_i
$\vec{\theta}$ is our given conditions, including β, ϕ, ω, and the effective population size
$\sum_j$ is shorthand for $\sum_{j=1}^{n}$, the sum over the whole ??gene/genome??
β_χ is the cost of codon translation for codon χ. In the NSE model, this is constant, $c_1 = c_2 = \cdots = c_n$. In the ROC model, this is the main variable.
ω_χ is the odds of a nonsense error for codon χ, $\frac{p_\chi}{1-p_\chi}$
ϕ is the expression level of the gene
$\mathbb{C}$ is the set of all possible ??synonymous?? sequences
ν is the number of synonymous codons to $\vec{c_i}$

3 Math

"The probability of seeing codon $\vec{c_i}$ at position i (given our conditions) is approximately the fitness of that sequence divided by the sum of the fitness of all possible synonymous sequences"

$$P(\vec{c_i}|\vec{\theta}) \approx \frac{\exp\left[-\sum_j \beta_{c_j}\omega_{c_j}(\vec{c_j})\phi\right]}{\sum_{K\in\mathbb{C}} \exp\left[-\sum_j \beta_{K_j}\omega_{K_j}(\vec{c_j})\phi\right]}$$

Rewrite the sum as the sum across the synonyms. Added notation:
ν is the number of synonyms for the given codon i

$$= \frac{\sum_{\kappa \in \mathbb{C} | \kappa_i = c_i} \exp\left[-\sum_j \beta_{\kappa_j} \omega_{\kappa_j}(\vec{\kappa}_j)\phi\right]}{\sum_{\ell=1}^{\nu} \sum_{K \in \mathbb{C} | K_{i\ell} = c_{i\ell}} \exp\left[-\sum_j \beta_{K_j} \omega_{K_j}(\vec{c}_j)\phi\right]}$$

Pull out the term for the current codon in the chain, where $c_j = c_i$

$$= \frac{\left[\sum_{\kappa \in \mathbb{C}} \exp\left[-\sum_{j \neq i} \beta_{\kappa_j} \omega_{\kappa_j}(\vec{\kappa}_j)\phi\right]\right] \left(\exp\left[-\beta_{c_i} \omega_{c_i}(\vec{c}_i)\phi\right]\right)}{\sum_{\ell=1}^{\nu} \sum_{K \in \mathbb{C} | K_{i\ell} = c_{i\ell}} \exp\left[-\sum_j \beta_{K_j} \omega_{K_j}(\vec{K}_j)\phi\right]}$$

Additional expansion that wasn't on the board.

$$= \frac{\left[\sum_{\kappa \in \mathbb{C}} \exp\left[-\sum_{j \neq i} \beta_{\kappa_j} \omega_{\kappa_j}(\vec{\kappa}_j)\phi\right]\right] \left(\exp\left[-\beta_{c_i} \omega_{c_i}(\vec{c}_i)\phi\right]\right)}{\sum_{\ell=1}^{\nu} \sum_{K \in \mathbb{C} | K_{i\ell} = c_{i\ell}} \exp\left[-\sum_{j \neq i} \beta_{K_j} \omega_{K_j}(\vec{K}_j)\phi\right] \exp\left[-\beta_{K_i\ell} \omega_{K_i\ell}(\vec{K}_i\ell)\phi\right]}$$

The ℓ term in the denominator can be pulled out of the inner sum, leaving

$$= \frac{\left[\sum_{\kappa \in \mathbb{C}} \exp\left[-\sum_{j \neq i} \beta_{\kappa_j} \omega_{\kappa_j}(\vec{\kappa}_j)\phi\right]\right] \left(\exp\left[-\beta_{c_i} \omega_{c_i}(\vec{c}_i)\phi\right]\right)}{\sum_{\ell=1}^{\nu} \left(\exp\left[-\beta_{K_i\ell} \omega_{K_i\ell}(\vec{K}_i\ell)\phi\right] \sum_{K \in \mathbb{C} | K_{i\ell} = c_{i\ell}} \exp\left[-\sum_{j \neq i} \beta_{K_j} \omega_{K_j}(\vec{K}_j)\phi\right]\right)}$$

Since the inner sum no longer relies on ℓ, we can restate the above as

$$= \frac{\left[\sum_{\kappa \in \mathbb{C}} \exp\left[-\sum_{j \neq i} \beta_{\kappa_j} \omega_{\kappa_j}(\vec{\kappa}_j)\phi\right]\right] \left(\exp\left[-\beta_{c_i} \omega_{c_i}(\vec{c}_i)\phi\right]\right)}{\sum_{\ell=1}^{\nu} \left(\exp\left[-\beta_{K_i\ell} \omega_{K_i\ell}(\vec{K}_i\ell)\phi\right]\right) \left(\sum_{K \in \mathbb{C}} \exp\left[-\sum_{j \neq i} \beta_{K_j} \omega_{K_j}(\vec{K}_j)\phi\right]\right)}$$

Which cancels to

$$P(\vec{c}_i | \vec{\theta}) \approx \frac{\exp\left[-\beta_{c_i} \omega_{c_i}(\vec{c}_i)\phi\right]}{\sum_{\ell=1}^{\nu} \exp\left[-\beta_{K_j} \omega_{K_j}(\vec{K}_j)\phi\right]}$$