We try to explain how the mixing rule of the second virial coefficients is obtained by analyzing the properties of binary mixtures statistically. The partition function in a grand canonical ensemble of binary gaseous mixtures reads
\[ \Xi = 1 + (\lambda_1 + \lambda_2) V + \frac{1}{2} \left( \lambda_1^2 \tilde{B}_{11} + 2 \lambda_1 \lambda_2 \tilde{B}_{12} + \lambda_2^2 \tilde{B}_{22} \right) V + \cdots \]
where \( \lambda_1 \equiv e^{\mu_1 / (k_B T) / \Lambda_1^3}, \lambda_2 \equiv e^{\mu_2 / (k_B T) / \Lambda_2^3} \), and
\[ \tilde{B}_{ij} \equiv \int \text{d} \mathbf{r} e^{-u_{i,j}(|\mathbf{r}|) / (k_B T)}, \quad (i, j = 1, 2) \]
\[ = \int \text{d} \mathbf{r} \left( e^{-u_{i,j}(|\mathbf{r}|) / (k_B T)} - 1 + 1 \right) = -\int \text{d} \mathbf{r} \left( 1 - e^{-u_{i,j}(|\mathbf{r}|) / (k_B T)} \right) + V \]
\[ = V - 2 B_{ij} \]
\[ B_{ij} \equiv \frac{1}{2} \int \text{d} \mathbf{r} \left( 1 - e^{-u_{i,j}(|\mathbf{r}|) / (k_B T)} \right) = 2\pi \int_0^\infty \text{d} r r^2 \left( 1 - e^{-u_{i,j}(r) / (k_B T)} \right) \]
Here \( \mu_i \) is the chemical potential of \( i \)-th component and \( \Lambda_i \) is the thermal de Broglie wavelength of the \( i \)-th component. For an ideal gas, the chemical potential is given by \( \mu_i = k_B T \ln(\rho_i / \Lambda_i) \), so \( \lambda_i \) reduces to the density \( \rho_i = N_i / V \). For non-ideal gas we are interested, \( \lambda_i \) play the role similar to \( \rho_i \) and are treated as small, because the virial expansion is most useful only when the density is low.

Then we can take the Taylor expansion for \( \ln \Xi \), in \( \lambda_i \), and keep the terms up to the quadratic order, which leads us to
\[ \ln \Xi = 1 + (\lambda_1 + \lambda_2) V + \frac{1}{2} \left( \lambda_1^2 \tilde{B}_{11} + 2 \lambda_1 \lambda_2 \tilde{B}_{12} + \lambda_2^2 \tilde{B}_{22} \right) V - \frac{1}{2} (\lambda_1 + \lambda_2)^2 V^2 + \cdots \]
Note that the constants \( V \) in \( B_{ij} \) cancel the last term, we get
\[ \ln \Xi = (\lambda_1 + \lambda_2) V - (\lambda_1^2 \tilde{B}_{11} + 2 \lambda_1 \lambda_2 \tilde{B}_{12} + \lambda_2^2 \tilde{B}_{22}) V \]
Since the pressure relates to the partition function through \( PV = k_B T \ln \Xi \), we then get
\[ \frac{p}{k_B T} = (\lambda_1 + \lambda_2) - (\lambda_1^2 \tilde{B}_{11} + 2 \lambda_1 \lambda_2 \tilde{B}_{12} + \lambda_2^2 \tilde{B}_{22}) \]
In order to get the virial expansion, we need to relate \( \lambda_i \) to density \( \rho_i \), to the same order. For the component-1, we get
\[ \rho_1 = \frac{N_1}{V} = \frac{1}{V} \frac{\partial \ln \Xi}{\partial \mu_1} = \frac{\lambda_1}{V} \frac{\partial \ln \Xi}{\partial \lambda_1} = \lambda_1 \left[ 1 - 2 (\lambda_1 \tilde{B}_{11} + \lambda_2 \tilde{B}_{12}) \right] \]
\[ \rho_2 = \lambda_2 \left[ 1 - 2 (\lambda_2 \tilde{B}_{22} + \lambda_1 \tilde{B}_{12}) \right] \]
and equivalently (up to the same order)
\[ \lambda_1 = \rho_1 + 2 \rho_1 (\rho_1 \tilde{B}_{11} + \rho_2 \tilde{B}_{12}) \]
\[ \lambda_2 = \rho_2 + 2 \rho_2 (\rho_2 \tilde{B}_{22} + \rho_1 \tilde{B}_{12}) \]
Substituting the above expressions into that for the pressure, we get (up to the same order of magnitude)
\[ \frac{p}{k_B T} = (\rho_1 + \rho_2) + (\rho_1^2 \tilde{B}_{11} + 2 \rho_1 \rho_2 \tilde{B}_{12} + \rho_2^2 \tilde{B}_{22}) \]
Noticing that the total density is \( \rho = \rho_1 + \rho_2 \), we further have
\[ \frac{p}{k_B T} = \rho + (y_1^2 \tilde{B}_{11} + 2 y_1 y_2 \tilde{B}_{12} + y_2^2 \tilde{B}_{22}) \rho^2 = \rho + B \rho^2 , \]
in which \( y_i \equiv N_i / (N_1 + N_2) = \rho_1 / \rho \), and the mixing rule for the second virial coefficient is found to be
\[ B = y_1^2 \tilde{B}_{11} + 2 y_1 y_2 \tilde{B}_{12} + y_2^2 \tilde{B}_{22} . \]
This is the desired result, which is exact even when many-body interactions are present.