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 $\lambda_1 \equiv e^{\mu_1/(k_{\mathrm{B}}T)}/\Lambda_1^3$, $\lambda_2 \equiv e^{\mu_2/(k_{\mathrm{B}}T)}/\Lambda_2^3$, and

$$\begin{aligned}
\tilde{B}_{ij} &\equiv \int \mathrm{d}\mathbf{r}\, e^{-u_{ij}(|\mathbf{r}|)/(k_{\mathrm{B}}T)}, \quad (i,j=1,2) \\
&= \int \mathrm{d}\mathbf{r}\, \left(e^{-u_{ij}(|\mathbf{r}|)/(k_{\mathrm{B}}T)} - 1 + 1\right) = -\int \mathrm{d}\mathbf{r}\, \left(1 - e^{-u_{ij}(|\mathbf{r}|)/(k_{\mathrm{B}}T)}\right) + V \\
&= V - 2B_{ij} \\
B_{ij} &\equiv \frac{1}{2}\int \mathrm{d}\mathbf{r}\, \left(1 - e^{-u_{ij}(|\mathbf{r}|)/(k_{\mathrm{B}}T)}\right) = 2\pi \int_0^\infty \mathrm{d}r\, r^2 \left(1 - e^{-u_{ij}(r)/(k_{\mathrm{B}}T)}\right)
\end{aligned}$$

Here μ_i is the chemical potential of i-th component and Λ_i is the thermal de Broglie wavelength of the i-th component. For an ideal gas, the chemical potential is given by $\mu_i = k_{\mathrm{B}}T \ln(\rho_i \Lambda_i)$, so λ_i reduces to the density $\rho_i = N_i/V$. For non-ideal gas we are interestedin, λ_i play the role similar to ρ_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 λ_i, and keep the terms up to the quadratic order, which leads us to

$$\ln \Xi = 1 + (\lambda_1 + \lambda_2)V + + \frac{1}{2}\left(\lambda_1^2 \tilde{B}_{11} + 2\lambda_1\lambda_2 \tilde{B}_{12} + \lambda_2^2 \tilde{B}_{22}\right) V - \frac{1}{2}(\lambda_2 + \lambda_2)^2 V^2 + \cdots$$

Note that the costants V in $\tilde{B}_{ij}$ cancel the last term, we get

$$\ln \Xi = (\lambda_1 + \lambda_2)V - \left(\lambda_1^2 B_{11} + 2\lambda_1\lambda_2 B_{12} + \lambda_2^2 B_{22}\right) V$$

Since the pressure relates to the partition function through $PV = k_{\mathrm{B}}T \ln \Xi$, we then get

$$\frac{p}{k_{\mathrm{B}}T} = (\lambda_1 + \lambda_2) - \left(\lambda_1^2 B_{11} + 2\lambda_1\lambda_2 B_{12} + \lambda_2^2 B_{22}\right)$$

In order to get the virial expansion, we need to relate λ_i to density ρ_i, to the same order. For the component-1, we get

$$\begin{aligned}
\rho_1 &= \frac{N_1}{V} = \frac{1}{V}\frac{\partial \ln \Xi}{\partial \mu_1} = \frac{\lambda_1}{V}\frac{\partial \ln \Xi}{\partial \lambda_1} = \lambda_1 \left[1 - 2\left(\lambda_1 B_{11} + \lambda_2 B_{12}\right)\right] \\
\rho_2 &= \lambda_2 \left[1 - 2\left(\lambda_2 B_{22} + \lambda_1 B_{12}\right)\right]
\end{aligned}$$

and equivalently (up to the same order)

$$\begin{aligned}
\lambda_1 &= \rho_1 + 2\rho_1(\rho_1 B_{11} + \rho_2 B_{12}) \\
\lambda_2 &= \rho_2 + 2\rho_2(\rho_2 B_{22} + \rho_1 B_{12})
\end{aligned}$$

Substituting the above expressions into that for the pressure, we get (up to the same order of magnitude)

$$\frac{p}{k_{\mathrm{B}}T} = (\rho_1 + \rho_2) + \left(\rho_1^2 B_{11} + 2\rho_1\rho_2 B_{12} + \rho_2^2 B_{22}\right)$$

Noticing that the total density is $\rho = \rho_1 + \rho_2$, we furthre have

$$\frac{p}{k_{\mathrm{B}}T} = \rho + \left(y_1^2 B_{11} + 2y_1y_2 B_{12} + y_2^2 B_{22}\right)\rho^2 = \rho + B\rho^2,$$

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

$$B = y_1^2 B_{11} + 2y_1y_2 B_{12} + y_2^2 B_{22}.$$

This is the desired result, which is exact even when many-body interactions are present.