The optimization problem is:

$$\min \frac{P_{TX}}{\eta(P_{TX})} + N P_{\text{circuit}} + P_{\text{Shared}}$$

where $\eta(P_{TX}) = a P_{TX} + b$ since the efficiency is modeled as a linear function of the transmit power.

If the transmit power under omni is $P_O$, we have $P_{TX} = P_O / N$ due to the beamforming gain.

Then the optimization problem (for $N$) is

$$\min \frac{P_O/N}{a(P_O/N) + b} + N P_{\text{circuit}} + P_{\text{Shared}}.$$

It is easy to solve and

$$N_{opt} = \frac{P_O}{\sqrt{C_1 P_{\text{circuit}}}} \frac{C_2 P_O}{C_1}$$

where $C_1 = b^2$ and $C_4 = a^2 b$ are constants.