Formal Justification:

1. Let $P_i = EX_i^2$ be the input power of the ith channel. Consider

$$\begin{aligned}
I(\mathbf{X};\mathbf{Y}) &= h(\mathbf{Y}) - h(\mathbf{Z}) \\
&\le \sum_{i=1}^{k} h(Y_i) - \sum_{i=1}^{k} h(Z_i) && (1) \\
&\le \frac{1}{2}\sum_{i=1}^{k} \log[2\pi e(EY_i^2)] - \frac{1}{2}\sum_{i=1}^{k} \log(2\pi e N_i) && (2) \\
&= \frac{1}{2}\sum_{i=1}^{k} \log(EY_i^2) - \frac{1}{2}\sum_{i=1}^{k} \log N_i \\
&= \frac{1}{2}\sum_{i=1}^{k} \log(EX_i^2 + EZ_i^2) - \frac{1}{2}\sum_{i=1}^{k} \log N_i \\
&= \frac{1}{2}\sum_{i=1}^{k} \log(P_i + N_i) - \frac{1}{2}\sum_{i=1}^{k} \log N_i && (3) \\
&= \frac{1}{2}\sum_{i=1}^{k} \log\left(1 + \frac{P_i}{N_i}\right).
\end{aligned}$$

2. The inequalities in (1) and (2) are tight when X_i's are independent and $X_i \sim \mathcal{N}(0, P_i)$.

3. Therefore, maximizing $I(\mathbf{X};\mathbf{Y})$ becomes maximizing $\sum_i \log(P_i + N_i)$ in (3).

4. The capacity of the system of parallel Gaussian channels is equal to the sum of the capacities of the individual Gaussian channels with the input power optimally allocated.