estimating the term $e(Y_1 \setminus Z_1, Y_2, \ldots, Y_k)$, we use random subsets $R_2, \ldots, R_k$ of size $(1-\varepsilon)s$ of $Y_2, \ldots, Y_k$, respectively. Thus,

$$e(Y_1 \setminus Z_1, Y_2, \ldots, Y_k) = (1-\varepsilon)s^k \mathbf{E}[\rho(Y_1 \setminus Z_1, R_2, \ldots, R_k)]\,.$$

Now for any choice of $R_2, \ldots, R_k$, we have

$$\begin{aligned}\rho(Y_1 \setminus Z_1, R_2, \ldots, R_k) &= ((1-\varepsilon)s)^{-\varepsilon^k} \mu(Y_1 \setminus Z_1, R_2, \ldots, R_k)\\ &\le ((1-\varepsilon)s)^{-\varepsilon^k} \mu(Y_1, Y_2, \ldots, Y_k)\\ &= (1-\varepsilon)^{-\varepsilon^k} \rho(Y_1, \ldots, Y_k).\end{aligned}$$

Therefore,

$$e(Y_1 \setminus Z_1, Y_2, \ldots, Y_k) \le (1-\varepsilon)^{1-\varepsilon^k} e(Y_1, \ldots, Y_k)$$

To estimate the term $e(Z_1, Z_2, \ldots, Z_{i-1}, Y_i \setminus Z_i, Y_{i+1}, \ldots, Y_k)$, we use random subsets $R_i \subset Y_i \setminus Z_i$ and $R_{i+1} \subset Y_{i+1}, \ldots,\ R_k \subset Y_k$, this time all of size εs. A similar calculation as before yields

$$e(Z_1, Z_2, \ldots, Z_{i-1}, Y_i \setminus Z_i, Y_{i+1}, \ldots, Y_k) \le \varepsilon^{i-1-\varepsilon^k}(1-\varepsilon) e(Y_1, \ldots, Y_k).$$

(This estimate is also valid for $i = 1$, but it is worse than the one derived above and it would not suffice in the subsequent calculation.) From (9.2) we obtain that $e(Z_1, \ldots, Z_k)$ is at least $e(Y_1, \ldots, Y_k)$ multiplied by the factor

$$\begin{aligned}1-(1-\varepsilon)^{1-\varepsilon^k} - (1-\varepsilon)\varepsilon^{-\varepsilon^k}\sum_{i=2}^{k} \varepsilon^{i-1} &= 1-(1-\varepsilon)^{1-\varepsilon^k} - \varepsilon^{1-\varepsilon^k} + \varepsilon^{k-\varepsilon^k}\\ &\ge 1-(1-\varepsilon)^{1-\varepsilon^k} - \varepsilon^{1-\varepsilon^k} + \varepsilon^k\\ &\ge 1-2^{\varepsilon^k} + \varepsilon^k\end{aligned}$$

where the last inequality follows from the inequality $\left(\frac{a^\alpha + b^\alpha}{2}\right)^{1/\alpha} \le \frac{a+b}{2}$, $a > 0$, $b > 0$, $0 < \alpha \le 1$, between the αth degree mean and the arithmetic mean. Now the function $f(x) = 1 - 2^x - x$ safisfies $f(0) = f(1) = 1$, and it is concave on $(0,1)$ since $f''(x) = -(\ln 2)^2 2^x < 0$. Hence $1 - 2^{\varepsilon^k} + \varepsilon^k > 0$ for all $\varepsilon \in (0,1)$ and Theorem 9.4.1 is proved. □

Bibliography and remarks. Our presentation of Theorem 9.4.1 essentially follows Pach [Pac98], whose treatment is an adaptation of an approach of Komlós and Sós.

The Szemerédi regularity lemma is from [Sze78], and in its full glory it goes as follows: *For every $\varepsilon > 0$ and for every k_0, there exist K and n_0 such that every graph G on $n \ge n_0$ vertices has a partition $(V_0, V_1, \ldots, V_k)$ of the vertex set into $k{+}1$ parts, $k_0 \le k \le K$, where $|V_0| \le \varepsilon n$, $|V_1| = |V_2| = \cdots = |V_k| = m$, and all but at most εk^2*