Fixing the integral argument in the Szeméredi Regularity Lemma

Let's recall some of the notation we had before the integral argument: We had

1. $\epsilon > 0$ fixed,
2. counter-examples to the Lemma G_K for each K,
3. an ultraproduct $G = \prod_{\mathcal{U}} G_K$ in the inductive language we created,
4. a measure on G we were calling μ that was the ultralimit of the counting measure on the G_K's and
5. we had formulas $U_1, \ldots, U_n$ in our language $\mathcal{L}$ which created a partition of G (and G_K for almost all K).

Each U_i, when interpreted, could be assumed to have measure greater than 0. This partition was related to the function h I described in the lecture in the following way:

$$h = \sum_{1 \leq i,j \leq n} \alpha_{i,j} \chi_{U_i} \times \chi_{U_j} + h'$$

where $\|h'\|_2 < \frac{\epsilon^4}{4}$ and $\|\chi_E - h\|_2 = 0$.

We wanted to try to show that $U_1, \ldots, U_n$ is ϵ-regular in an appropriate sense in G. Towards this end, we were computing the measure of the set B of bad pairs i, j; that is, the pairs for which

$$R_{U_i,U_j} and S_{U_i,u_j} \text{ are not empty for ultrafilter many } K$$

or equivalently, $R_{U_i,U_j} and S_{U_i,u_j}$ are non-empty in G. For such a bad i, j, we let

$$\beta_{i,j} = \lim_{\mathcal{U}} \frac{|E \cap R_{U_i,U_j} \times S_{U_i,u_j}|}{|R_{U_i,U_j}||S_{U_i,u_j}|} = \frac{\mu(E \cap R_{U_i,U_j} \times S_{U_i,u_j})}{\mu(S_{U_i,u_j})\mu(R_{U_i,U_j})}$$

For any i, j, I want to compute $d(U_i, U_j)$ i.e. the edge density measured via μ between these two sets. We have

$$\int \chi_E \chi_{U_i \times U_j} d\mu = \mu(E \cap (U_i \times U_j)$$

and by Cauchy-Schwartz

$$\int (\chi_E - h)\chi_{U_i \times U_j} = 0$$

so after rearranging we get

$$\mu(E \cap (U_i \times U_j) = \alpha_{i,j}\mu(U_i)\mu(U_j) + \int h' \chi_{U_i \times U_j} d\mu.$$

The last term is over-estimated by $\frac{\epsilon^4}{4}\mu(U_i)\mu(U_j)$ and when we divide by $\mu(U_i)\mu(U_j)$ we have

$$d(U_i, U_j) \leq \alpha_{i,j} + \frac{\epsilon^4}{4}.$$

For bad i, j we also know that $|d(U_i, U_j) - \beta_{i,j}| \geq \epsilon$ and so putting this altogether, we have

$$|\beta_{i,j} - \alpha_{i,j}| \geq \epsilon - \frac{\epsilon^4}{4} \text{ which we will call } \delta.$$

Note that by possibly choosing ϵ small enough, we can assume that $\delta \geq \frac{\epsilon}{2}$.

Now suppose that B^+ is the set of i, j in B for which $\alpha_{i,j} \geq \beta_{i,j} + \delta$ and

$$Z = \cup_{i,j \in B^+} R_{U_i, U_j} \times S_{U_i, U_j}.$$

Toward a contradiction, suppose that $\mu(\cup_{i,j \in B^+} U_i \times U_j) \geq \frac{\epsilon}{2}$. Then, using a similar argument as above, we have

$$\begin{aligned}
\left| \int h' \chi_Z d\mu \right| &= \left| \int \left(\sum_{i,j} \alpha_{i,j} \chi_{U_i \times U_j} \chi_Z - \chi_E \chi_Z \right) d\mu \right| \\
&= \left| \sum_{i,j \in B^+} \left(\alpha_{i,j} \mu(R_{U_i, U_j}) \mu(S_{U_i, U_j}) - \int \chi_E \chi_{R_{U_i, U_j} \times S_{U_j, U_j}} d\mu \right) \right| \\
&\geq \left| \sum_{i,j \in B^+} \left(\alpha_{i,j} \mu(R_{U_i, U_j}) \mu(S_{U_i, U_j}) - (\alpha_{i,j} - \delta) \mu(R_{U_i, U_j}) \mu(S_{U_i, U_j}) \right) d\mu \right| \\
&= \sum_{i,j \in B^+} \delta \mu(R_{U_i, U_j}) \mu(S_{U_i, U_j})
\end{aligned}$$

and since $\mu(R_{U_i,U_j})$ and $\mu(S_{U_i,U_j})$ are greater than $\epsilon\mu(U_i)$ and $\epsilon\mu(U_j)$ respectively, this latter sum is greater than $\delta\epsilon^2\frac{\epsilon}{2}$ which in turn is greater than $\frac{\epsilon^4}{4}$. But again by Cauchy-Schwartz, $|\int h'\chi_Z| < \frac{\epsilon^4}{4}$ which is a contradiction. So we conclude that $\mu(\cup_{i,j\in B^+}U_i \times U_j) < \frac{\epsilon}{2}$. A very similar argument gives us that $\mu(\cup_{i,j\in B\setminus B^+}U_i \times U_j) < \frac{\epsilon}{2}$. We conclude then that $\mu(\cup_{i,j\in B}U_i \times U_j) < \epsilon$.

This shows that the partition $U_1, \ldots, U_n$ is ϵ-regular in the sense of the measure μ. By Łoś, for almost all K this is true in G_K and when $K > n$, this contradicts the original choice of G_K.