Lemma 2. There exists a $G$-invariant network in the sense of definition 3 that realizes the sum of $G$-invariant networks $F = \sum_{k=0}^{d} \sum_{j=1}^{n_k} \alpha_{kj} F^{kj}$.

Proof. We need to show that $F = \sum_{k=0}^{d} \sum_{j=1}^{n_k} \alpha_{kj} F^{kj}$ can indeed be realized as a single, unified $G$-invariant network. As we already saw, each network $F^{kj}$ has the structure

$$
R^n \xrightarrow{L} R^{k \times k} \xrightarrow{M} R^n \xrightarrow{U} R^n
$$

with a suitable $k$-class $\tau$. To create the unified $G$-invariant network we first lift each $F^{kj}$ to the maximal dimension $d$. That is, $\tilde{F}^{kj}$ with the structure

$$
R^n \xrightarrow{\tilde{L}^{kj}} R^{d \times d} \xrightarrow{\tilde{M}^{kj}} R^n \xrightarrow{\tilde{U}} R^n
$$

This is done by composing each equivariant layer $L : R^n \times a \to R^n \times b$ with two linear equivariant operators $U^b : R^n \times b \to R^{d \times b}$ and $D^a : R^n \times a \to R^{n \times a}$,

$$
U^b LD^a : R^{d \times a} \to R^{n \times b},
$$

where

$$
U^b(x)_{i_1...i_d,j} = x_{i_1...i_d,j}
$$

and

$$
D^a(y)_{i_1...i_k,j} = n^{d-k} \sum_{i_{k+1}...i_d=1} y_{i_1...i_{k+1}...i_d,j}.
$$

Since $U^b, D^a$ are equivariant, $U^b LD^a$ in equation 1 is equivariant. Furthermore $D^a \circ \sigma \circ U^a = \sigma$, where $\sigma$ is the pointwise activation function. Lastly, given two $G$-invariant networks with the same tensor order $d$ they can be combined to a single $G$-invariant network by concatenating their features. That is, if $L_1 : R^{n \times a} \to R^{n \times b}$, and $L_2 : R^{n \times a'} \to R^{n \times b'}$, then their concatenation would yield $L_{1,2} : R^{n \times (a + a')} \to R^{n \times (b + b')}$. Applying this concatenation to all $\tilde{F}^{kj}$ we get our unified $G$-invariant network. \hfill \Box

Fixed-point equation for equivariant layers. We have an affine operator $L : R^n \times a \to R^n \times b$ satisfying

$$
g^{-1} \cdot L(g \cdot X) = L(X),
$$

for all $g \in G, X \in R^n \times a$. The purely linear part of $L$ can be written using a tensor $L \in R^{n \times a \times b}$: Write

$$
L(X)_{j_1...j_d,j} = \sum_{i_1...i_k} L_{j_1...j_i...j_k,i,j} X_{i_1...i_k,i}.
$$

Writing equation 2 using this notation gives:

$$
\sum_{i_1...i_k,i} L_{g(j_1)...g(j_i),i_1...i_k,i,j} X_{g^{-1}(j_1)...g^{-1}(i_k),i}
= \sum_{i_1...i_k,i} L_{g(j_1)...g(j_i),i_1...i_k,i,j} X_{i_1...i_k,i}
= \sum_{i_1...i_k,i} L_{j_1...j_i...j_k,i} X_{i_1...i_k,i},
$$

for all $g \in G$ and $X \in R^n \times a$. This implies equation 9, namely

$$
g \cdot L = L, \quad g \in G.
$$

The constant part of $L$ is done similarly.

\begin{flushright}
1Department of Computer Science and Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel. 2Department of Computer Science, University of Toronto, Toronto, Canada. Correspondence to: Haggai Maron <haggai.maron@weizmann.ac.il>.
\end{flushright}