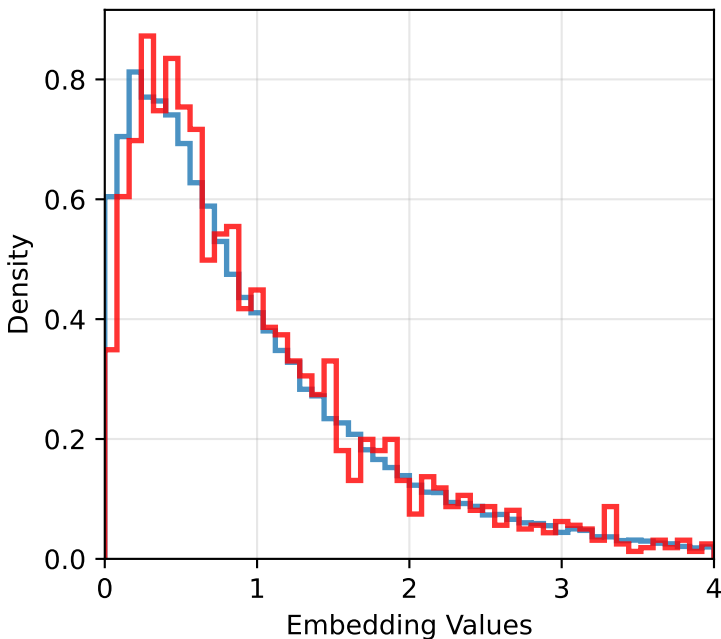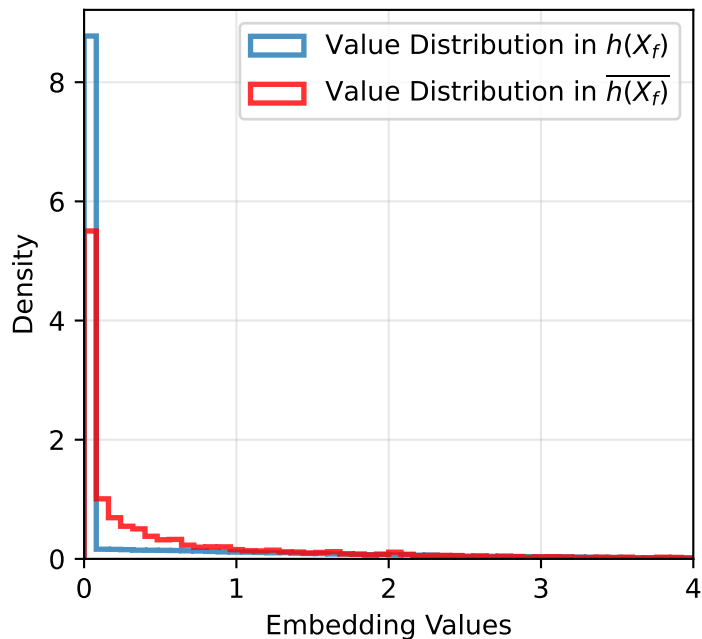Distribution Shift for ReLu before vs after averaging (variance normalized)