7.4 **Pretty good measurements.** Consider a density operator ρ and a pure-state ensemble decomposition

$$\rho = \sum_{\alpha=1}^{N} p_\alpha |\psi_\alpha\rangle\langle\psi_\alpha| = \sum_\alpha |\overline{\psi}_\alpha\rangle\langle\overline{\psi}_\alpha| \ ,$$

where $|\overline{\psi}_\alpha\rangle = \sqrt{p_\alpha}|\psi_\alpha\rangle$. The support of ρ is the subspace spanned by the eigenvectors with nonzero eigenvalues. The HJW theorem for ensemble decompositions tells us that the decomposition vectors $|\overline{\psi}_\alpha\rangle$ lie in and span the support. In this problem, we restrict attention to the support, which we assume to be D-dimensional (with $D \le N$), and forget about the rest of Hilbert space. This means that ρ is invertible.

We can define a POVM that has POVM elements

$$E_\alpha = \rho^{-1/2}|\overline{\psi}_\alpha\rangle\langle\overline{\psi}_\alpha|\rho^{-1/2} = |\overline{\phi}_\alpha\rangle\langle\overline{\phi}_\alpha| \ ,$$

where $|\overline{\phi}_\alpha\rangle = \rho^{-1/2}|\overline{\psi}_\alpha\rangle = \sqrt{p_\alpha}\rho^{-1/2}|\psi_\alpha\rangle$. The POVM elements are clearly positive operators and the POVM satisfies the completeness relation

$$\sum_\alpha E_\alpha = \rho^{-1/2} \underbrace{\left(\sum_\alpha |\overline{\psi}_\alpha\rangle\langle\overline{\psi}_\alpha|\right)}_{=\rho} \rho^{-1/2} = I \ .$$

This measurement is called the *pretty good measurement.*

(a) *Show* that the outcome probabilities $q_\alpha = \mathrm{tr}(\rho E_\alpha)$, when the state is ρ, are the same as the ensemble probabilities p_α. This is the unique property of the pretty good measurement.

(b) *Show* that if the decomposition vectors $|\overline{\psi}_\alpha\rangle$ are linearly independent, there thus being $N = D$ of them, $\langle\overline{\phi}_\alpha|\overline{\phi}_\beta\rangle = \delta_{\alpha\beta}$, i.e., that the POVM is an ODOP. For equal probabilities p_α, this is a nice way to generate an orthonormal basis from a set of linearly independent vectors, which in contrast to the Gram-Schmidt process, doesn't favor any vector in the set.

(c) For the case $N = D = 2$ and equal prior probabilities, *show* that the pretty good measurement is the measurement that minimizes error probability.