Suppose we have the causal model consisting of U, A, X with domain $\mathbb{U}$, $\mathbb{X}$ and $\mathbb{A}$. The prior distribution of U and A are encoded in the probability density functions $p(U)$ and $p(A)$.

The joint distribution of U, A, X can be written as

$$p(U, A, X) = p(U)p(A)\delta(X - f(A, U)) \tag{4}$$

From Bayes theorem, we know

$$p(U|A, X) = \frac{p(U, A, X)}{p(X, A)} = \frac{p(U)p(A)\delta(X - f(A, U))}{\int_{\mathbb{U}} p(U)p(A)\delta(X - f(A, U))dU} \tag{5}$$

With the same definition of H, we can know

$$p(H|X, A) = \delta\left(\int_{\mathbb{U}} \frac{p(U)p(A)\delta(X - f(U, A))}{\int_{\mathbb{U}} p(U)p(A)\delta(X - f(U, A))dU}[s(\cdots), U] - H\right) \tag{6}$$

The conditional distribution of H on A is

$$p(H|A) = \int_{\mathbb{X}} p(H|X, A)p(X|A)dX \tag{7}$$

And

$$p(X|A) = \frac{P(X, A)}{p(A)} = \frac{\int_{\mathbb{U}} p(A)p(U)\delta(X - f(A, U))dU}{\int_{\mathbb{U}} \int_{\mathbb{X}} p(A)p(U)\delta(X - f(A, U))dUdX} \tag{8}$$

which equals to

$$p(X|A) = \frac{P(X, A)}{p(A)} = \frac{\int_{\mathbb{U}} p(A)p(U)\delta(X - f(A, U))dU}{p(A)} \tag{9}$$

As a result,

$$p(H|A) = \int_{\mathbb{X}} \delta\left(\int_{\mathbb{U}} \frac{p(U)p(A)\delta(X - f(U, A))}{\int_{\mathbb{U}} p(U)p(A)\delta(X - f(U, A))dU}[s(\cdots), U] - H\right) \frac{\int_{\mathbb{U}} p(A)p(U)\delta(X - f(A, U))dU}{p(A)} dX \tag{10}$$

Because U follows a uniform distribution, the probability density function can be simplified as

$$p(H|A) = \int_{\mathbb{X}} \delta\left(\int_{\mathbb{U}} \frac{\delta(X - f(U, A))}{\int_{\mathbb{U}} \delta(X - f(U, A))}[s[\cdots], U] - H\right) \left(\int_{\mathbb{U}} p(U)\delta(X - f(U, A))dU\right) dX \tag{11}$$

The condition 3 which requires

$$f(u_1, a) = f(u_2, a) \Leftrightarrow f(u_1, a') = f(u_2, a') \tag{12}$$

makes us know for every u in $\mathbb{U}$, it belongs to a sub-space $\mathbb{U}^s \subset \mathbb{U}$. The sub-space $\mathbb{U}^s$ satisfies

$$f(u_1, a) = f(u_2, a) \quad u_1, u_2 \in \mathbb{U}^s, a \in \mathbb{A} \tag{13}$$

We assume the number of these sub-spaces is countable, namely $\mathbb{U}^1$, $\mathbb{U}^2, \cdots$, $\mathbb{U}^k \cdots$. We denote

$$\int_{\mathbb{U}^k} 1dU = \Omega^k \tag{14}$$

Since $P(U)$ is a uniform distribution,

$$\int_{\mathbb{U}} p(U)dU = 1 \tag{15}$$

Denote $\{1_X, \cdots, k_X, \cdots\}$ as the sub-spaces where $f(U,A) = X$, we have

$$\int_{\mathbb{U}} p(U)\delta(X - f(U,A))dU = \frac{\sum \Omega^{1_X} + \cdots \Omega^{k_X} + \cdots}{\sum \Omega^1 + + \cdots \Omega^k + \cdots} \tag{16}$$

$$\int_{\mathbb{U}} \delta(X - f(U,A))dU = \sum \Omega^{1_X} + \cdots \Omega^{k_X} + \cdots \tag{17}$$

$$\int_{\mathbb{U}} \delta(X - f(U,A))[s(\cdots), U]dU = \sum \int_{\mathbb{U}^{k_X}} c_{k_X} dU = \sum c_{k_X} \Omega^{k_X} \tag{18}$$

in which c_{k_X} is a constant for each subspace $\mathbb{U}^{k_X}$ because X and $\check{X}$ are constants inside each of the subspace.

So,

$$p(H|A) = \int_{\mathbb{X}} \delta(\frac{\sum c_{k_X} \Omega^{k_X}}{\sum \Omega^{k_X}} - H)\frac{\sum \Omega^{1_X} + \cdots \Omega^{k_X} + \cdots}{\sum \Omega^1 + + \cdots \Omega^k + \cdots} dX \tag{19}$$

For any specific value $X \in \mathbb{X}$, we can have $\check{X} \in \mathbb{X}$. Since $X = f(U,A)$ and $\check{X} = f(U,\breve{A})$, we know $\{1_X, \cdots, k_X, \cdots\}$ are the sub-spaces where $f(U,\breve{A}) = \check{X}$.

Because $c_{k_X} = c_{k_{\check{X}}}$, we know that

$$p(H|\breve{A}) = \int_{\check{\mathbb{X}}} \delta(\frac{\sum c_{k_{\check{X}}} \Omega^{k_X}}{\sum \Omega^{k_X}} - H)\frac{\sum \Omega^{1_X} + \cdots \Omega^{k_X} + \cdots}{\sum \Omega^1 + + \cdots \Omega^k + \cdots} dX = p(H|A) \tag{20}$$

which is to say $H \perp A$.