Here is an idea for a generalisation of channels with memory.

Let $\mathcal{M}$ be a finite-dimensional C*-algebra. (This is the memory space.) Suppose that $E : \mathcal{B}(\mathcal{H}) \otimes \mathcal{M} \otimes \mathcal{M} \to \mathcal{B}(\mathcal{H}) \otimes \mathcal{M}$ is a conditional expectation, i.e. a completely positive map with $E^2 = E$, if we consider $\mathcal{B}(\mathcal{H}) \otimes \mathcal{M}$ as a subspace of $\mathcal{B}(\mathcal{H}) \otimes \mathcal{M} \otimes \mathcal{M}$. If we embed $\mathcal{B}(\mathcal{H}) \otimes \mathcal{M}$ into

$$\mathfrak{A} = \overline{\cup_n \mathfrak{A}_n}, \text{ with } \mathfrak{A}_n = (\mathcal{B}(\mathcal{H}) \otimes \mathcal{M})^{\otimes n}, \tag{1}$$

then we write $E_{k,k-1}$ if we apply E to the $(k-1)$-th and k-th factors, i.e.

$$\begin{aligned} E_{k,k-1} \left(\sum_{j_1,\dots,j_n} \lambda_{j_1,\dots,j_n} A_{j_1} \otimes \cdots \otimes A_{j_n} \otimes B_{j_1} \otimes \cdots \otimes B_{j_n} \right) = \\ = \sum_{j_1,\dots,j_n} \lambda_{j_1,\dots,j_n} A_{j_1} \otimes B_{j_1} \otimes \cdots \otimes A_{j_{k-2}} \otimes B_{j_{k-2}} \otimes A_{j_{k-1}} \otimes \\ E(A_{j_k} \otimes B_{j_{k-1}} \otimes B_{j_k}) \otimes \cdots \otimes B_{j_n}. \end{aligned} \tag{2}$$

Let ρ be a state on $\mathcal{M}$ such that the conditional expectation $E_\rho : \mathcal{B}(\mathcal{H}) \otimes \mathcal{M} \to \mathcal{B}(\mathcal{H})$ satisfies

$$E_\rho(E(A \otimes B)) = E_{\rho\otimes\rho}(A \otimes B) \tag{3}$$

for $A \in \mathcal{B}(\mathcal{H})$ and $B \in \mathcal{M} \otimes \mathcal{M}$. Given a state ϕ on

$$\mathfrak{B} = \overline{\cup_n \mathcal{B}(\mathcal{H}^{\otimes n})}, \tag{4}$$

we can extend it to a state on $\mathfrak{A}$ via

$$\tilde{\phi}^{(n)}(A \otimes B) = \phi^{(n)}(A) \rho^{\otimes n}(B) \tag{5}$$

for $A \in \mathcal{B}(\mathcal{H}^{\otimes n})$ and $B \in \mathcal{M}^{\otimes n}$.

Now we can define a completely positive map $\Phi : \mathcal{S}(\mathfrak{B}) \to \mathcal{S}(\mathfrak{B})$ by

$$\Phi^{(n)}(\phi^{(n)})(A) = \tilde{\rho}^{(n)} \left[E_{2,1} \circ \cdots \circ E_{n,n-1}(A \otimes \mathbf{1}) \right]. \tag{6}$$

The classical memory example fits into this scheme as follows: We take $\mathcal{M} = \mathbb{C}^M$ with pointwise product, i.e. $(v_i)_{i=1}^M \cdot (w_i)_{i=1}^M = (v_i w_i)_{i=1}^M$. This is the same as considering $\mathbb{C}^M$ as the algebra of continuous functions on $\{1, \dots, M\}$. An element $A \in \mathcal{B}(\mathcal{H}) \otimes \mathbb{C}^M$ can be written uniquely as

$$A = \sum_{j=1}^M A_j \otimes e_j, \tag{7}$$

where $\{e_j\}_{j=1}^{M}$ is the standard basis of $\mathbb{C}^M$ and $A_j \in \mathcal{B}(\mathcal{H})$. Then we define E by

$$E\left(\sum_{j_1,j_2=1}^{M} A_{j_1,j_2} \otimes e_{j_1} \otimes e_{j_2}\right) = \sum_{j_1,j_2=1}^{M} q_{j_2|j_1} \left[(1 \otimes V_{j_2}^*) A_{j_1,j_2} (\mathbf{1} \otimes V_{j_2})\right] \otimes e_{j_1}. \tag{8}$$

In order to get ergodicity, we need conditions on the conditional expectation E, which are not clear to me at the moment. I did check in Hugenholtz (AMS, Proc. Symp. Pure Math. **38** (1982), Part 2): the weak mixing condition is what we need for ergodicity:

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=0}^{m-1} \phi(A\, g_k(B)) = \phi(A)\, \phi(B). \tag{9}$$

In the classical memory case, this is easy to prove. In fact, we have strong mixing:

$$\lim_{m\to\infty} \phi(A\, g_k(B)) = \phi(A)\, \phi(B). \tag{10}$$