We formalize the intuition that the state space of dynamical systems encodes the relevant information in the past system input about the system's future output. In that sense, the state space serves as *sufficient statistic* from the system's past to future. We employ the Information Bottleneck (IB) method [1] which is a general information theoretic solution to the problem of extracting compressed representations of random variables that while preserving information about other variables. We apply this method in order to infer state-spaces that maximize the predictive information [2] about the output of the system. Here we specifically apply this new formalism to linear dynamical systems with Gaussian state and input noise, relying on our previous work [3] that solved the information bottleneck problem for Gaussian variables. In this paper we provide a general solution to the question: what kind of dynamical system optimally captures the relevant information of the input past about the output future through a state space bottleneck. We state the IB problem through a Lagrangian for dynamical systems and formulate the relevant mutual information quantities in terms of the covariance matrices which explicitly dependent on the state space. This allows us to unravel the input history in the IB sense in terms of structural phase transitions corresponding to additional dimensions of the state space, or poles and zeros of the model system transfer function. The critical values for these transitions are functions of the well known Hankel singular values in control problems [4]. It reveals known system identification algorithms in this new context [4-6]. This and related work [7] explicitly link information theory with the general framework of dynamical system inference through the important notion of predictive information.

[1] N. Tishby, F. Pereira, and W. Bialek, in *Proc. of 37th Allerton Conference on Communication and Computation (1999)*.

[2] W. Bialek, I. Nemenman, N. Tishby, *Neural Computation* 13, 2409-2463 (2001)

[3] G. Chechik, A. Globerson, N. Tishby, and Y. Weiss, *Journal of Machine Learning Research (JLMR)* 6, 165-188 (2005)

[4] S. Gugercini and A. C. Antoulas, *Int.J.Control* 77, 748 (2004).

[5] C. Mullis and R. Roberts, *IEEE Trans. Circuits Syst*. CAS-23, 551 (1976).

[6] B.C.Moore, IEEE Trans. *Automat. Contr*. AC-26, 17 (1981).

[7] F. Creutzig, A. Globerson, and T. Tishby, Tech. Rep., Computer Science, Hebrew University (2005).