mation. When we measure the new k_S value for the same nodes in the resulting networks (Fig. 6) we find that their relative ranking remains the same. We recover a practically linear dependence on the k_S values of the original and the incomplete networks, showing that this measure would work equally well for predicting the spreading efficiency of nodes in a network with missing information.

III. PROBABILITY AND TIME OF INFECTION

We have demonstrated that the location of a node, as described through the k_S index, is important for the extent of spreading M_i when this node is the spreading origin. Here, we show that nodes with high k_S are more probable to be infected during an epidemic outbreak and are infected earlier than nodes with low k_S, when spreading starts at a random node. We introduce the quantity E_i, as the probability that a node i is going to be infected during an epidemic outbreak originating at a random location, and T_i, as the average time before node i is infected during the same process.

As shown in Figs. 7a-d all three quantities that characterize the role of a node in an epidemics process, M_i, E_i and T_i are strongly correlated. The nodes that are infected by a given node i form a cluster of size $\overline{M_i}$, and they are statistically the nodes that can reach i when they act as origins themselves. Thus, the probability E_i to reach this node in general is directly proportional to the size M_i, as shown in the plots. The average time T_i to reach a node is inversely proportional to its spreading efficiency M_i, which emphasizes the fact that these nodes are easily reachable from different network locations. In conclusion, the nodes with the largest k_S values consistently a) are infecting larger parts of the network, b) are infected more frequently, and c) are infected earlier, than nodes with smaller k_S values.

IV. THE IMPRECISION FUNCTIONS

We quantify the spreading efficiency of an individual origin i through the infected number of nodes M_i. In order to compare the different methods, we rank all network nodes according to their spreading efficiency, independently of their other properties, and we consider a fraction p of the most efficient spreaders ($p \in [0,1]$). We designate this set by $\Upsilon_{eff}(p)$. Similarly, we define $\Upsilon_{k_S}(p)$ as the set of individuals with highest k-shell values. In order to