In this paper, the downlink nonorthogonal multiple access (NOMA) system is studied where purely discrete input distributions are found that achieve the capacity region to within a constant gap without successive interference cancellation (SIC). The approach is a two-step approach where the corresponding linear deterministic model is first studied and the results are then systematically translated into purely discrete input distributions for the original model. A simple yet powerful coding scheme, which adopts off-the-shelf turbo codes with pulse amplitude modulations (PAM) is then used to simulate the proposed input distributions. Simulation results show that the proposed simple scheme under turbo decoding, both with and without SIC, can operate close to information-theoretic bounds of the proposed input distributions, which lies outside the achievable rate region of any orthogonal multiple access (OMA)-type scheme.