

ACOUSTICS2008/2455
**A nonlinear frequency-domain beamformer for underdetermined
speech mixtures**

Michael Davies^a and Mohammad Dmour^b

^aUniversity of Edinburgh, ICDOM, Kings Buildings, Mayfield Road, EH9 3JL Edinburgh, UK

^bUniversity of Edinburgh, Institute for Digital Communications, School of Engineering & Electronics,
Alexander Graham Bell Building, Kings Buildings, Mayfield Road, EH9 3JL Edinburgh, UK

Extraction of a target speech source from among multiple interfering speech sources is challenging when there are fewer microphones than sources (the underdetermined case). Existing speech source separation techniques often suffer from artifacts as well as performance deterioration in reverberant environments, and in some cases also need to estimate the number of sources present. This paper introduces a frequency-domain non-linear beamformer that can perform speech source separation of underdetermined mixtures, is reasonably artifact free and does not require prior knowledge of the number of speakers. Our approach models the data via a Gaussian mixture distribution in the observation domain, which can be learned using the expectation maximization (EM) algorithm. A non-linear distortionless beamformer is then developed, based on this model. Simulations of the non-linear beamformer in underdetermined mixtures with room reverberation confirm its capability to successfully separate speech sources with virtually no artifacts.