

# Combining the multicanonical ensemble with generative probabilistic models of local biomolecular structure

Jes Frellsen<sup>1,2,†</sup>, Thomas Hamelryck<sup>2</sup>, Jesper Ferkinghoff-Borg<sup>3</sup>

<sup>1</sup>Department of Engineering, University of Cambridge, United Kingdom

<sup>2</sup>Department of Biology, University of Copenhagen, Denmark

<sup>3</sup>Department of Systems Biology, Technical University of Denmark, Denmark

<sup>†</sup>Corresponding author: Jes Frellsen, e-mail: jf519@cam.ac.uk

## Abstract

Markov chain Monte Carlo is a powerful tool for sampling complex systems such as large biomolecular structures. However, the standard Metropolis-Hastings algorithm suffers from a number of deficiencies when applied to systems with rugged free-energy landscapes. Some of these deficiencies can be addressed with the multicanonical ensemble. In this paper we will present two strategies for applying the multicanonical ensemble to distributions constructed from generative probabilistic models of local biomolecular structure. In particular, we will describe how to use the multicanonical ensemble efficiently in conjunction with the reference ratio method.

Keywords: Markov chain Monte Carlo, multicanonical ensemble, generative probabilistic models, biomolecular structure.