

High resolution timing from model based deconvolution of timecourse microarray data with maximum entropy prior

Andrzej Kudlicki, Maga Rowicka and Zbyszek Otwinowski

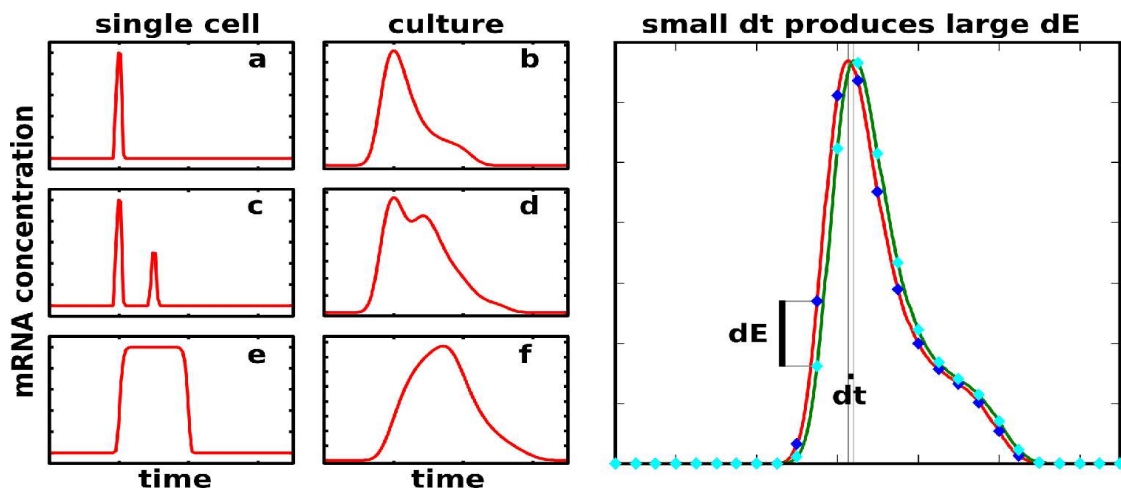
University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75390-8816, USA

In a timecourse microarray experiment, whole-genome expression data are collected along the time line of a sequential or periodic biological process. The measured mRNA concentrations are averages over cells in the culture, therefore the temporal resolution in such experiments is limited by both data sampling and cell to cell synchronization. We show how imperfect synchronization of the cell culture can be actually turned into an advantage, provided that an approximate model of cell synchrony is known. The measured transcript concentration, M , can be viewed as a convolution of the single cell expression profile S and the population spread function (or asynchrony model) P :

$$M(t) = \int_{-\infty}^{\infty} S(u)P(t-u)du$$

We present a solution for $S(u)$ given the measurements $M(t_i)$, the population model P , and a model of measurement errors. Our method is based on conjugate gradient optimization, where the target function consists of a goodness of the fit term, and a term corresponding to the entropy of the distribution. The information contained in the population model allows to reconstruct the single cell profile with a temporal resolution an order of magnitude higher than the data sampling. It works very well with noisy data, because noise not following the asynchrony model is automatically filtered out.

Our procedure has been successfully applied to analysis of cell cycle in yeast and human cell cultures, yielding temporal maps of transcriptionally regulated events of an unprecedented accuracy. Other potential applications include circadian rhythms and development processes. The nature of source data is not limited to mRNA levels: temporal profiles of metabolite concentrations, or FACS counts may also be deconvolved.



Left: examples of single cell expression profiles [a,c,e] and corresponding culture average [b,d,f]. Typical cell-cycle asynchrony model is shown. Right: Small differences in moment of expression peak produce significant changes in measured expression level; thus allowing for high temporal resolution.