IEF Final Report (327999 CosmicDawn21cm)

Geraint Harker (Fellow), Filipe Abdalla (Scientist in charge)

1st March 2014 – 29th Feb. 2016

1 Introduction

The Cosmic Dawn is the period between approximately 400 000 and 1 billion years after the Big Bang. At its start, the Universe was relatively uniform and featureless, but by the end it was populated by stars, galaxies, black holes and other astrophysical objects. The radiation from the first stars and galaxies has a profound influence on the rest of the matter in the Universe, heating and ionizing it. The latter part of this period is therefore referred to as the Epoch of Reionization, and is the last major phase transition in the Universe's history.

The effect of the first stars on the surrounding medium allows this period to be studied with low-frequency radio telescopes. Most of the matter in the Universe is in the form of hydrogen which, unless ionized, emits and absorbs radiation at a characteristic wavelength of 21 cm. This emission or absorption occurs when the spin axis of the single electron in a hydrogen atom flips its direction, changing between being in the same direction as that of the nucleus and the opposite direction. 21-cm radiation originating from the Cosmic Dawn and the Epoch of Reionization is stretched by the expansion of Universe, such that by the time it reaches us it is at wavelengths larger than 1 m, corresponding to frequencies of between tens of kilohertz and around 200 MHz. The frequency at which it is detected can be converted directly into a measure of how much the Universe has expanded since it set off, and therefore how old the Universe was at the time. Meanwhile, the intensity of the radiation contains information about the state of the Universe at that time, though it requires further modelling and interpretation to relate this to the properties of the first stars and galaxies.

Several experiments are being conducted with the aim of detecting this radiation, and more are planned or are under construction. Though the techniques developed in this project can, for the most part, be applied more generally, the work was directed towards two of these in particular: *DARE*, a proposed lunar-orbiting satellite that would probe the middle of the Cosmic Dawn, and *LOFAR*, a radio telescope operating in the Netherlands which aims to study the epoch of reionization. *LOFAR* is also a pathfinder for the (much larger) Square Kilometre Array, the low-frequency part of which will be built in Australia, and will be used to study the Cosmic Dawn in more detail.

2 Contribution of this project

Signals from the Cosmic Dawn reach us at low radio frequencies, at which a large number of other astrophysical objects also produce radiation. These 'foregrounds' pose a difficult signal separation problem since they are much more intense than the Cosmic Dawn signal, and this problem requires several steps to overcome. Firstly, the foregrounds and the signal from the Cosmic Dawn must be modelled so that we have some prior constraints on the signal separation. Secondly, the measurement process itself must be understood, since it distorts and filters the signals from the sky. Thirdly, we must develop suitable statistical methods which have the potential to perform the signal separation under these conditions. Fourthly, the analysis must be simulated and tested, using synthetic data which are as realistic as possible, so that the signal separation pipeline can be refined, and so that it can be demonstrated that a detection of the Cosmic Dawn signal is possible. This also informs the design of the experiments to make the detection. Finally, the pipeline is applied to the experimental data so that we can make inferences about the Cosmic Dawn.

This project has contributed in two main areas: developing simple models of the Cosmic Dawn signal which are suitable for use in the statistical inference of the properties of the first stars; and developing analysis pipelines which can be used on experimental data, and have been applied to synthetic data to demonstrate the feasibility of Cosmic Dawn experiments.

2.1 Analysis techniques

The precise form of the foregrounds for sky-averaged measurements of the 21-cm line from the Cosmic Dawn and the Epoch of Reionization is unknown. Before and during the project we developed a pipeline to perform Bayesian inference of the parameters of foreground models, simultaneously with estimation of parameters of models of the instrument and 21-cm signal, using Monte Carlo techniques.

More specifically, over the period of the project, we went further, showing that the level of complexity in the foreground models should itself be driven by the data, under a Bayesian model selection methodology. We applied a technique known as nested sampling to simplified models of global 21-cm data to compute the Bayesian evidence for the models. Some physical arguments and preliminary measurements suggest that the foregrounds might usefully be modelled as polynomials of order n in log-log space. If $n = 4$ provides a more accurate model than $n = 3$, this can be inferred with less than two hours of integration with and experiment with limited frequency coverage, but that otherwise resembles *DARE*, for reasonable values of the $n = 4$ coefficient. If a higher-order polynomial is required, this does not necessarily prevent a significant detection of the 21-cm signal. Even for $n = 8$, which groups working on other experiments have argued may be more realistic, we can obtain very strong evidence distinguishing a reasonable model for the signal from a null model with 128 h of integration. More subtle features of the signal may, however, be lost if the foregrounds are this complex. This was demonstrated using a simpler model for the signal that only includes absorption (as expected during the Cosmic Dawn) and neglecting emission. We highlighted some pitfalls in trying to quantify the significance of a detection from errors on the parameters of the signal alone.

2.2 Models of the 21-cm signal

Following our previous work, which related generic features in the sky-averaged (global) 21-cm signal to properties of the intergalactic medium, we investigated the prospects for constraining a simple galaxy formation model with current and near-future experiments. We developed fast, sophisticated codes to generate plausible models of the signal, and combined this with realistic models for foregrounds from our Galaxy, noise consistent with 100 h of integration with a well-behaved instrument, and a simple model for the other distortions introduced by this instrument.

We then performed Markov-Chain Monte Carlo fits to the resulting synthetic data set, finding that a simple four-parameter model can be well-constrained (to around 25% in each dimension, orders of magnitude better than what has been achieved so far) so long as three distinctive features in the 21-cm spectrum, which are expected to occur between 40 and 120 MHz, are indeed detected. The model links the production rate of photons with different energy bands (which are important in governing the temperature and ionization state of the Universe's hydrogen, and how this state affects the properties of the 21-cm radiation produced by the gas) to the growth rate of dark matter haloes.

Several important conclusions turn out to follow naturally from this basic numerical result, namely that measurements of the global 21-cm signal can in principle (i) identify the mass threshold for dark matter haloes during the Cosmic Dawn, above which formation of the first stars can proceed inside these haloes, (ii) extend our knowledge about the relation between the quantity of X-rays produced by stars and the star formation rate to much earlier times, informing us about the properties of these first stars, and (iii) provide joint constraints on stellar spectra and on how much of the ionizing radiation produced by stars during the Epoch of Reionization can escape the neighbourhood of the stars and affect the surrounding medium. Though our approach was general, we show the necessity of performing the measurement over the broadest frequency range possible and of having a clean view of the global 21-cm signal; this renders our findings most relevant to the *DARE*, which would have the advantage of observing above the radio-quiet, ionosphere-free lunar far-side.