Bayesian Analysis of Excitonic Spectra in Cu_2O*

High energy exciton peaks become hard to distinguish because their oscillator strengths decrease as n^{-3} and their spacings decrease as n^{-2}, causing the peaks to overlap as they approach the bandgap [1]. Because of this, least squares fitting algorithms are not adequate for discerning high energy exciton resonances (this is why we studied only resonances with $n \leq 7$ in the previous sections). A more effective fitting can be done with Bayesian reconstruction [2], which helps to identify even those features which have intensities below the noise level.

The fundamental premise of Bayesian analysis is determining what underlying "model" is the most likely to have generated the data measured in an experiment, given that the system has some underlying random noise profile. It has far-reaching applications in image processing, where faint images can be extracted from noisy backgrounds such as deep space or biological tissue [3–5]. It has also been used to study the temporal evolution of a diverse range of systems from galaxies to plasma clouds and reconstruct their prior states [6, 7]. Here, we employ this technique to extract fine details about high-energy Rydberg exciton states with intensities below the noise level.

1 Theory of Bayesian Analysis

For a given dataset, D, with n data points given by x-values x_i and y-values y_i that is being fit with some function, f, with fit parameters given by the set θ, the raw dataset can be described as

$$y_i = f(x_i;\theta) + N(0,\sigma)\,, \tag{1}$$

where $N_i(\sigma)$ represents the Gaussian background noise overlaid on the dataset, which has a mean of 0 and a standard deviation of σ. Using Bayes' Law, we can ascertain that

$$P(\theta|D) \propto P(D|\theta)P(\theta)\,, \tag{2}$$

where $P(\theta|D)$ represents the probability of a given set of fit parameters given the measured dataset, $P(\theta)$ is the prior probability of the fit parameters, and $P(D|\theta)$ is the probability of a dataset being measured given an underlying set of fit parameters, described by the equation

$$P(D|\theta) \propto \exp\big[-\frac{n\varepsilon(\theta)}{\sigma^2}\big]\,, \tag{3}$$

where $\varepsilon(\theta)$ is the mean-squared error (MSE) between the dataset and the model. In the context of curve fitting, this allows us to gain information about the probability distribution of the fit parameters, $P(\theta|D)$, by evaluating how well various parameter sets describe the dataset.

To leverage this, the *Metropolis-Hastings algorithm*, as described in [8], can be used to sample from $P(D|\theta)$ and estimate the underlying distribution of the fit parameters. Its procedure is as fllows:

1. At the start of the algorithm, a set of guess parameters, θ_0, is proposed alongside a set of guess standard deviations for each parameter, σ_θ.
2. A candidate parameter set, θ^*, is randomly generated as

$$\theta^* = \theta_i + N(0,\sigma_\theta)\,. \tag{4}$$

3. A probability ratio r is then calculated as

$$r = \frac{P(\theta^*|D)}{P(\theta_i|D)}\,. \tag{5}$$

*This guide has been adapted from chapter 4 of DeLange, Jacob C (2023). Optical Properties of Rydberg Excitons in Cuprous Oxide. Purdue University Graduate School. Thesis. https://doi.org/10.25394/PGS.22590112.v1

4. If $r > 1$, indicating that θ^* fits the dataset better than θ_i, the candidate set is accepted unconditionally ($\theta^* \to \theta_{i+1}$). If $r < 1$, indicating that θ_i fits better than θ^*, a random number u is sampled uniformly from the range [0,1]. If $u < r$, the candidate set is still accepted, but if $u > r$, it is rejected ($\theta_i \to \theta_{i+1}$).

5. Steps 2-4 are then repeated for a set number of iterations which must be chosen sufficiently large to allow the algorithm to converge to the true underlying distribution of the fit parameters.

Once the algorithm has been run, the stored parameter sets θ_i can be used to approximate the fit parameter's true probability distribution, from which the means and variances of the parameters can be estimated. However, this method requires prior knowledge of a system's underlying noise level, which is not realistic. Furthermore, it provides no benchmark by which to compare different models, which is essential for our purpose of determining how many peaks are present in a spectrum. To resolve this issue, we use the *Replica Exchange Monte Carlo (RXMC) algorithm*. This protocol generalizes the Metropolis-Hastings algorithm to test multiple potential noise levels simultaneously. At the start of the algorithm, a separate set of guess parameters is initialized for each noise value being tested. These "replicas" are then iterated according to the following steps:

1. An iteration of the metropolis algorithm is conducted for each replica.

2. Next, an opportunity is given for neighboring replicas (corresponding to similar background noise values) to exchange their paremter sets. For replicas at positions i and $i+1$, a parameter s is calculated as

$$s = \frac{P(\theta_{i+1}|D,f,\sigma_i)P(\theta_i|D,f,\sigma_{i+1})}{P(\theta_i|D,f,\sigma_i)P(\theta_{i+1}|D,f,\sigma_{i+1})}. \tag{6}$$

3. If $s > 1$, indicating that the fits for the two replicas would be improved by an exchange, then the exchange is always permitted ($\theta_i \to \theta_{i+1}$, $\theta_{i+1} \to \theta_i$). If $s < 1$, indicating that the fits for the two replicas would not be improved by an exchange, then a random number v is sampled uniformly from the range [0,1]. If $v < s$, the exchange is permitted. If $v > s$, the exchange is forbidden.

4. Steps 2 and 3 are then repeated sufficiently many times to allow the distribution to be appropriately sampled. At this point, each replica will have a list of associated parameter sets which can be used to estimate the probability distribution of the fit parameters for the background noise value associated with each replica.

To determine the correct value of background noise inherent to the dataset, we consider for each replica a value Z, which is a function of both the model, f, and the background noise of that replica [2, 9]

$$Z(f,\sigma_i) = P(D|f,b) = \int d\theta P(D|\theta,f,\sigma_i)P(\theta|f,\sigma_i). \tag{7}$$

Put into words, Z represents a weighted average over all possible fit parameters of the probability of the measured dataset being generated by an underlying model and background noise. The weight function, $P(\theta|f,\sigma_i)$, can be estimated from the output of the RXMC algorithm. If the probabilities are thought of as "enegies," then Z is seen to be analogous with the partition function of a thermodynamic system. From this, a "free energy" can be calculated as

$$F(f,\sigma_i) = -\log Z(f,\sigma_i). \tag{8}$$

In practical calculations, F is used instead of Z because it can be conveniently approximated as

$$F(f,\sigma_i) = \frac{-1}{b_i}\sum_{j=1}^{i-1} \log\langle \exp[-n(b_{j+1}-b_j)\varepsilon(\theta_j)]\rangle, \tag{9}$$

where $b_i = \sigma_i^{-2}$, and the angular brackets represent an average over all parameter sets in the j^{th} replica.

As with statistical mechanics, where the most probable configuration minimizes the Helmholtz free energy, F is minimized here by the most probable value of the background noise. Using this information, the means and variances of the parameter set can be estimated using the data from the most probable replica. This approach can also be used to compare different types of models which might be used to fit the data. For example, if we repeated the process above with a fitting function g which achieved a lower minimum free energy, we would conclude that g is more likely to be the underlying model than f. In our case, the different models we test correspond to different numbers of peaks. When their minimum values of F are compared, we ascertain what number of Rydberg exciton peaks are most likely to be present given some measured dataset.

2 Bayesian Fitting of Excitonic Spectra from Cu_2O

Depending on measurement type from which the analyzed data originates, different lineshapes are used to fit the excitonic resonances. For an optical density measurement, asymmetric Fano lineshapes [10] are used,

$$\alpha_n(E) = f_n \frac{\frac{\Gamma_n}{2} + 2q_n(E - E_n)}{\left(\frac{\Gamma_n}{2}\right)^2 + (E - E_n)^2}, \tag{10}$$

whereas Lorentzian profiles are used for photoluminescence (PL) measurements [11]. Additionally, optical density measurements must be fit with an Urbach tail background, arising from continuum absorption which increases at higher energies. The Urbach tail is a heuristic model which signifies an increase in absorption near the bandgap energy caused by an increasing density of states in that energy range. This increase is described by an exponential function,

$$\alpha_U(E) = \alpha_0 \exp\left(\frac{E - E_g}{E_u}\right), \tag{11}$$

where E_g is the bandgap energy, α_0 is the magnitude of the continuum absorption, and E_u is the Urbach energy [12, 13]. For PL measurements, the background counts can be modeled with a constant offset.

To implement the Bayesian analysis, multiple models (corresponding to different numbers of peaks) are fit to the same dataset using the RXMC method. The parameters of these models can be subdivided into peak-independent and peak-dependent parameters. The peak-independent parameters are the bandgap energy, E_g, and the Rydberg binding energy, Ry, the Urbach absorption coefficient, a_0, and the Urbach energy, E_u, and a constant offset, c, which model the exponential behavior near the bandgap energy with an Urbach tail. The peak-dependent parameters are the linewidths, Γ_n, the Fano asymmetry factors, q_n, the peak-heights, f_n, and the quantum defects, δ_n. The quantum defects are used in-tandem with the bandgap and binding energies to calculate the center energies of the resonances, and this center energy in turn is combined with the other three peak-dependent parameters to model each resonance with a Fano or Lorentzian lineshape.

References

1. Aßmann, M. & Bayer, M. Semiconductor Rydberg Physics. *Adv. Quantum Technol.* **3,** 1900134. https://onlinelibrary.wiley.com/doi/full/10.1002/qute.201900134 (11 2020).
2. Tokuda, S., Nagata, K. & Okada, M. Simultaneous estimation of noise variance and number of peaks in Bayesian spectral deconvolution. *ArXiv* **abs/1607.07590** (2016).
3. Pina, R. K. & Puetter, R. C. BAYESIAN IMAGE RECONSTRUCTION: THE PIXON AND OPTIMAL IMAGE MODELING. *Publications of the Astronomical Society of the Pacific* **105,** 630. https://dx.doi.org/10.1086/133207 (June 1993).
4. Alenius, S. & Ruotsalainen, U. Bayesian image reconstruction for emission tomography based on median root prior. *European Journal of Nuclear Medicine* **24,** 258–265. ISSN: 1619-7089. https://doi.org/10.1007/BF01728761 (Mar. 1997).
5. Qi, J., Leahy, R. M., Cherry, S. R., Chatziioannou, A. & Farquhar, T. H. High-resolution 3D Bayesian image reconstruction using the microPET small-animal scanner. *Physics in Medicine & Biology* **43,** 1001. https://dx.doi.org/10.1088/0031-9155/43/4/027 (Apr. 1998).
6. Dikovsky, M. *et al.* Multi-instrument Bayesian reconstruction of plasma shape evolution in the C-2W experiment. *Physics of Plasmas* **28,** 062503. https://doi.org/10.1063/5.0049530 (2021).
7. Valade, A., Libeskind, N. I., Hoffman, Y. & Pfeifer, S. Testing Bayesian reconstruction methods from peculiar velocities. *Monthly Notices of the Royal Astronomical Society* **519,** 2981–2994. ISSN: 0035-8711. https://doi.org/10.1093/mnras/stac3673 (Dec. 2022).
8. Iwamitsu, K., Aihara, S., Okada, M. & Akai, I. Bayesian Analysis of an Excitonic Absorption Spectrum in a Cu_2O Thin Film Sandwiched by Paired MgO Plates. *Journal of the Physical Society of Japan* **85,** 094716. https://doi.org/10.7566/JPSJ.85.094716 (2016).
9. Iwamitsu, K., Okada, M. & Akai, I. Spectral Decomposition of Components Weaker than Noise Intensity by Bayesian Spectroscopy. *Journal of the Physical Society of Japan* **89,** 104004. https://doi.org/10.7566/JPSJ.89.104004 (2020).

10. Fano, U. Effects of Configuration Interaction on Intensities and Phase Shifts. *Phys. Rev.* **124,** 1866–1878. https://link.aps.org/doi/10.1103/PhysRev.124.1866 (6 Dec. 1961).

11. Takahata, M. & Naka, N. Photoluminescence properties of the entire excitonic series in Cu_2O. *Phys. Rev. B* **98,** 195205. https://link.aps.org/doi/10.1103/PhysRevB.98.195205 (19 Nov. 2018).

12. Urbach, F. The Long-Wavelength Edge of Photographic Sensitivity and of the Electronic Absorption of Solids. *Phys. Rev.* **92,** 1324–1324. https://link.aps.org/doi/10.1103/PhysRev.92.1324 (5 Dec. 1953).

13. Cody, G. D., Tiedje, T., Abeles, B., Brooks, B. & Goldstein, Y. Disorder and the Optical-Absorption Edge of Hydrogenated Amorphous Silicon. *Phys. Rev. Lett.* **47,** 1480–1483. https://link.aps.org/doi/10.1103/PhysRevLett.47.1480 (20 Nov. 1981).