Inference of mean-variance relationship was performed using a dataset from BMDCs incorporating 29 scRNA-seq experiments (each corresponding to a single Fluidigm C1 experiment with up to 96 cells) on the response time-course (at 0, 1, 2, 3 and 6 h) as well as additional perturbations such as treatment with IFNβ, inhibition of paracrine secretion (chemical or physical on chip) or cell knockout for ifnar1 and STAT1 expression (Shalek et al., 2014). We considered 813 genes that were induced by at least two-fold (compared to unstimulated cells) at the population level at any time point during the LPS stimulation [as identified in (Shalek et al., 2014)]. Visual inspection of the data revealed outliers in the linear regression fit, therefore, outlier removal method with Mahalanobis distance was used (with *0.05* threshold for outlier detection) (Finch, 2012). After removing low abundant genes (maximum mean expression <100 read counts) this resulted in 290 genes for the core TLR dataset (time-course) and 323 for the combined dataset (including perturbations). Bursting characteristics (based on moment estimators) for individual data points were fitted using linear regression and power functions (in semi log scale) when appropriate and presented as smooth curves. Robust regression (excluding data points with corresponding fitted residuals > 1.5⋅σresiduals) was used to either remove noisy data (as expected in the scRNA-seq measurement) or remove individual data-point that did not affect the overall trend. Equations, fitted parameters, corresponding correlation coefficients and highlighted outliers are included in the encosed files. Fitting protocols were implemented in Python using R kernel, individual gene graphs were produced in MATLAB R2014a.

1. Folder ‘\_pre-processing’ includes:

* Calculation of summary statistics (mean, variance, moment estimators of burst size and frequency, etc) for all TLR-induced genes and datasets using ‘summary\_stats.ipynb’; raw data (read counts per gene per C1 chip) is stored in ‘induced\_gene\_raw\_data.csv’, summary statistics are stored in ‘induced\_genes\_processed\_data\_all\_genes\_all\_conditions.csv’.
* Application of Mahalonobis distance to (repeatedly) remove measurement outliers (per gene) using ‘mahalonobis\_1.ipynb’; output is stored in ‘induced\_genes\_outliers\_removed\_all\_812\_4runs.csv’

1. Folder ‘core TLR analysis’ includes regression fitting of mean-variance relationship for core genes in the dataset (TLR2, TL3 and TLR4 stimulation). Fitting is performed using ‘my\_mean\_var\_regression.ipynb’ using pre-processed data ‘induced\_genes\_outliers\_removed\_all\_812.csv’ (after removing conditions with zero counts). Code in Matlab (‘validation.m’) provides a detailed display of fitted data using robust regression.
2. Folder ‘core TLR with perturbation’ includes curve fitting of different characteristics in the full dataset. These involve applications of robust regression as well as comparison with the core TLR dataset for mean-variance, mean-burst-size, mean-frequency as well as burst size variance and frequency-variance. Each subfolder is organised in analogous way to ‘core TLR analysis’.