

Proceedings THE THERMODYNAMICS OF URBAN AND REGIONAL STRUCTURE

Louis Ellam ^{1,2}*, Mark Girolami ^{1,2}, Grigorios Pavliotis ¹ and Alan Wilson ^{2,3}

- ¹ Department of Mathematics, Imperial College London, London, SW7 2AZ
- ² The Alan Turing Institute, The British Library, London, NW1 2DB
- ³ Centre for Advanced Spatial Analysis, University College London, London, W1T 4TJ
- * Correspondence: l.ellam16@imperial.ac.uk

Version May 21, 2018 submitted to Proceedings

Abstract: It is well recognized that the mathematics underlying thermodynamics and statistical mechanics has wide applicability. In the context of urban modelling, entropy maximization has played an important role in deriving a family of spatial interaction models for flows in cities. Whilst urban and regional structure can be modelled by a dynamical system, much less has been recognised in connection with statistical mechanics. In this article, we develop a thermodynamic analogy and connect a stochastic reformulation of the Harris and Wilson model with a maximum entropy argument. That is, the probability distribution of the structural variables can be represented as a gradient flow of a free energy functional, and that the energy functional satisfies a second law of thermodynamics. We illustrate our model with the London retail system.

Keywords: Urban Modelling, Urban Structure, Complexity Science, Maximum Entropy, Energy Minimization, Fokker-Planck equation, Thermodynamics, Stochastic Differential Equations

1. Introduction

Thermodynamics and statistical mechanics provide valuable approaches for complex systems modelling, and the underlying techniques have wide applicability. The principle of maximum entropy has been used in a variety of domains, but has its origins in equilibrium thermodynamics. Its development was motivated by the task of relating the macroscopic properties of a physical system to behaviour at the atomistic level. Its wide applicability became apparent in the 1950s, when it was discovered that the tools of information theory can be applied to complex systems [1]. Today, the principle of maximum entropy is well-recognised by statisticians as a way of deriving probability distributions from the exponential family [2], and the connection with equilibrium thermodynamics and information theory provides a supporting heuristic.

For our purposes, there are two core elements of urban and regional systems: flows between location, which involve spatial interaction, and the urban and regional structure that facilitates the flows. The principle of maximum entropy has already been applied to the modelling of urban flows. In the 1960s, it was recognised that the proceeding so-called gravity model can be reformulated using the tools of the statistical mechanics [3]. The connection with statistical mechanics was developed by way of analogy, and it was subsequently recognized that entropy maximization provides a more powerful modelling framework [4]. A family of spatial interaction models has subsequently been derived from the framework for locational analysis [5].

Whilst entropy maximization has been applied with much success to the modelling of flows in cities, much less has been considered for urban or regional structure. Instead, non-linear dynamical systems were developed in the 1970s to model the evolution of structure [6], and models of the like have been used more recently in [7,8]. It has recently been acknowledged in the urban modelling

maximum entropy argument. It remains to fully develop an analogy between thermodynamics and the modelling of urban and regional systems that considers the non-equilibrium case. The connection between stochastic dynamical systems and the second law of thermodynamics have already been made in [12], and have not been considered in an urban context. The two schools of thought between dynamical systems and the principle of maximum entropy has also been connected in [13,14].

To this end, the contribution of this article is to establish a thermodynamic analogy for urban and regional structure. We first outline the principle of maximum entropy, and show how its application can be used to derive a model of urban and regional structure from the exponential family. We second show that a stochastic reformulation of the Harris and Wilson model satisfies a second law of thermodynamics, in that a free energy functional is monotonically decreasing until the corresponding Fokker-Planck equation converges to the maximum entropy distribution. We illustrate the analogy by studying the forward evolution of the London retail system [15]. We conclude with our research agenda and possibilities for developing the thermodynamic analogy further.

2. Maximum Entropy Models for Urban and Regional Structure

In this section, we present a maximum entropy model of urban structure. We assume that there are *M* competing urban zones of interest, and represent urban structure by the latent attractiveness variables $X = (X_1, \ldots, X_M)^T \in \mathbb{R}^M$. The attractiveness variables are defined as the log-sizes so that

$$X_j = \ln W_j. \tag{1}$$

We represent the structural variables in terms of attractiveness, rather than size, to develop a thermodynamic analogy. This way the stationary distribution of a related stochastic differential equation model, defined in terms of attractiveness, coincides with the maximum entropy distribution [11]. We first discuss the possible constraints for urban and regional systems, and then connect the maximum entropy distribution with the related dynamics by a thermodynamic analogy in the following section.

2.1. The Maximum Entropy Principle

The maximum entropy principle states that, amongst a family of probability distributions satisfying some constraints, we should choose the one that maximizes an entropy function [1]. We let $\mathcal{X} \subset \mathbb{R}^M$ be a finite domain with a smooth boundary $\partial \mathcal{X}$. A finite domain is appropriate for an urban system due to economic or geographical restrictions. We use the standard Boltzmann-Gibbs definition of entropy

$$S(\rho) = -\int_{\mathcal{X}} \rho(x) \ln \rho(x) dx < \infty,$$
(2)

where the integral is taken over the values in \mathcal{X} for which $\rho > 0$. Here, $\rho(x)$ denotes the probability density function of the structural variables X. The entropy function is bounded above as \mathcal{X} is a finite domain, and an upper bound can always be obtained from Gibbs inequality

$$-\int_{\mathcal{X}}\rho(x)\ln\rho(x)dx \leq -\int_{\mathcal{X}}\rho(x)\ln\rho'(x)dx,$$
(3)

with equality if and only if the two density functions ρ and ρ' coincide. We consider the case that the urban system is constrained by the expected potential energy

$$\langle V \rangle = \int_{\mathcal{X}} V(x) \rho(x) dx,$$
 (4)

where $V \in L^{\infty}(\mathcal{X})$ is a suitably chosen potential function for an urban system. The Boltzmann-Gibbs measure with the density function

$$\rho_*(x) = Z^{-1} e^{-\gamma V(x)}, \qquad Z = \int_{\mathcal{X}} e^{-\gamma V(x)} dx, \tag{5}$$

maximizes the Boltzmann-Gibbs entropy in (2) subject to the constraint in (4). The parameter γ is determined implicitly from the constraint from (4). The claim can be verified using Gibbs inequality in (3), since

$$H(\rho) \leq -\int_{\mathcal{X}} \rho(x) \ln \rho_*(x) dx,$$

= $H(\rho_*),$ (6)

with equality if and only if $\rho = \rho_*$. We now consider some possibilities for the potential function in the following subsection.

2.2. Constraints for Urban Systems

In this section we consider economic constraints for the distribution of the sizes of M urban zones, for example, shopping centres. To account for spatial interaction, we assume that the destination zones earn income from N origin zones. The origin zones are typically residential areas, and the origin quantities may be taken to be the spending power¹. A model of spatial interaction can be obtained by maximizing an entropy function subject to benefit and cost constraints. The reader is referred to [3,16] for further details. The spatial interaction model that we use is

$$D_{j}(X) = \sum_{i=1}^{N} O_{i} \frac{\exp(\alpha X_{j} - \beta c_{ij})}{\sum_{k=1}^{M} \exp(\alpha X_{k} - \beta c_{ik})}, \qquad j = 1, \dots, M.$$
(7)

where α , $\beta > 0$ are scaling parameters, and c_{ij} represents the inconvenience of carrying out an activity at zone *j* from *i*.

It is expected that profitable zones will reinvest profits to improve their attractiveness to consumers, whilst loss-making zones will their attractiveness to a more sustainable level. A suitable model that captures this evolutionary behaviour is a system of ordinary differential equations (ODE)s, known as the Harris and Wilson model [6]

$$\frac{dW_j}{dt} = \epsilon W_j (D_j - \kappa W_j), \qquad W_j(0) = w_j.$$
(8)

Here, W_j is the size of the respective zone and is connected to the attractiveness variables by (1). The $"D_j - \kappa W_j''$ term can be interpreted as the profitability of zone *j*, where we have assumed that the incomes D_j are given by the spatial interaction model in (7), and that $\kappa > 0$ describes the linear cost per unit size to run the zone. It can be shown that the ODE model will converge to a fixed point satisfying

$$D_i = \kappa e^{X_j}.\tag{9}$$

In fact, for $0 < \alpha \le 1$ there is a unique stable fixed point, otherwise there may be multiple fixed points [17].

Equation (9) suggests a constraint at equilibrium. In the same way that the spatial interaction model in (7) was obtained, we look towards the maximum entropy principle for a model of urban structure. One approach, already considered in the literature, is to maximize an entropy function

¹ Number of residents multiplied by the average earnings.

subject to a constraint on the total profitability [9,10]. This approach yields the following potential function

$$V(X) = \sum_{j=1}^{M} (D_j - \kappa e^{X_j}).$$
 (10)

As the total profit is equivalent to the total origin quantities less total capacity, the constraint is not restrictive enough as individual zones may be far from satisfying the equilibrium condition in (9). An alternative approach would be to constrain the total sum of squares, which gives a smooth potential function

$$V(X) = \sum_{j=1}^{M} (D_j - \kappa e^{X_j})^2.$$
 (11)

Whilst an improvement, it lacks economic justification and another shortfall is that unstable fixed points are assigned high probabilities. Our recent work in [11] unveils that a suitable constraint is the total consumer surplus (benefit) less total capacity (cost)

$$V(X) = -\alpha^{-1} \sum_{i=1}^{N} O_i \ln \sum_{j=1}^{M} \exp\left(\alpha X_j - \beta c_{ij}\right) + \kappa \sum_{j=1}^{M} e^{X_j}.$$
 (12)

The potential function assigns highest probabilities to the most stable configurations, as desired. The potential function and Harris and Wilson model are closely related as the minima of the potential function coincide with the fixed points of the Harris and Wilson model.

2.3. An Additional Constraint on the Attractiveness

Lastly, we specify an additional constraint on the first moment of attractiveness at equilibrium, following our previous work [11]. We therefore obtain the potential function

$$V(X) = -\alpha^{-1} \sum_{i=1}^{N} O_i \ln \sum_{j=1}^{M} \exp\left(\alpha X_j - \beta c_{ij}\right) + \kappa \sum_{j=1}^{M} e^{X_j} - \delta \sum_{j=1}^{M} X_j,$$
(13)

where the additional constraint is captured by the right-most term. Our initial justification was that the potential function needs to be confining on an unbounded domain. Whilst this is no longer necessary, as we are working on a bounded domain, the additional constraint provides an improved model. For instance, the constraint means that \mathcal{X} can be chosen to contain a local minima of V, and the model is less sensitive to the choice of boundary.

3. The Thermodynamic Analogy

Whilst the maximum entropy principle provides a way of obtaining a probability distribution for urban structure, there are some philosophical and practical arguments against its use. We may alternatively take a more mechanistic approach, and describe stochastic dynamics that has a well-defined equilibrium distribution. In the next section we describe a stochastic differential equation whose associated Fokker-Planck equation satisfies a second law of thermodynamics.

3.1. The Fokker-Planck equation

In this section we study the Fokker-Planck equation for a time-homogeneous process, which can be written in divergence form as

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot J,$$

$$J = -\nabla V \rho - \gamma^{-1} \nabla \rho,$$
(14)

for $x \in \mathcal{X}$. Here, $\rho(x, t)$ denotes the probability density of a process X(t) at time $t \ge 0$, and J is the probability flux. We assume that (14) is equipped with a reflecting boundary condition

$$J \cdot n = 0, \qquad x \in \partial \mathcal{X}, \tag{15}$$

so that the total probability is conserved in \mathcal{X} . It is well known that the stochastic differential equation (SDE) underlying (14) is

$$dX = -\nabla V(X)dt + \sqrt{2\gamma^{-1}}dB, \qquad X(0) = x_0,$$
(16)

with reflection at the boundaries, and where *B* is a standard *M*-dimensional Brownian motion. We refer the reader to [18] for a more mathematical description of reflecting SDEs. In other words, the Fokker-Planck equation is describing a process that experiences drift, diffusion and reflection. After having specified the potential function via a maximum entropy argument in the proceeding section, the SDE that we study is

$$dX_j = \epsilon \left[D_j - \kappa e^{X_j} + \delta \right] dt + \sqrt{2\gamma^{-1}} dB_j, \qquad j = 1, \dots, M.$$
(17)

This is a version of the Harris and Wilson model, with an additional constant in the drift function, and with a diffusion term. The diffusion term describing random fluctuations in order for the model to converge to the equilibrium distribution obtained under constrained entropy maximization. Otherwise the process will be deterministic and its long term behaviour is determined by the initial conditions.

3.2. The Trend To Equilibrium

We now show that (14) satisfies a second law of thermodynamics, and evolves towards the maximum entropy distribution given by (5). The relevant form of entropy of a second law of thermodynamics is proportional to a relative entropy, known as the free energy functional

$$F(\rho) = \int_{\mathcal{X}} V(x)\rho(x)dx + \gamma^{-1} \int_{\mathcal{X}} \rho(x)\ln\rho(x)dx.$$
 (18)

Here, γ has the interpretation as an inverse temperature. The free energy functional may be written as the difference of an energy functional and an entropy functional. It follows by Gibbs' inequality that

$$F(\rho) \ge \int_{\mathcal{X}} V(x)\rho(x)dx + \gamma^{-1} \int_{\mathcal{X}} \rho(x)\ln\rho_*(x)dx,$$

= $F(\rho_*),$ (19)

with equality if and only if $\rho = \rho_*$. Therefore the minimization of free energy yields the same probability distribution as constrained entropy maximization does. At equilibrium, the free energy functional becomes what is known as the Helmholtz free energy in the thermodynamics literature:

$$F(\rho_*) = \gamma^{-1} \ln Z. \tag{20}$$

We can reformulate the Fokker-Planck equation in terms of the free energy functional that we would like to minimize. We deduce that the evolution of $\rho(x, t)$ is a gradient flow of the free energy functional

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \frac{\delta F}{\delta \rho} \right), \qquad x \in \mathcal{X}.$$
(21)

Version May 21, 2018 submitted to Proceedings

Moreover, it can be shown that free energy functional is monotonically decreasing along the flow lines

$$\frac{d}{dt}F(\rho) = -\int_{\mathcal{X}} \rho \left| \nabla \frac{\delta F}{\delta \rho} \right|^2 dx, \qquad (22)$$
$$\leq 0,$$

which we show in Appendix. A. The result is the second law of thermodynamics for an open system. A key observation for our purposes is that the free energy is only decreasing when the system undergoes an irreversible change. For this reason, the free energy of the deterministic system with $\gamma \rightarrow \infty$ is as specified by the initial condition, and there is no thermodynamic analogy to be made. This is because the diffusion term vanishes in the Fokker-Planck equation, and it becomes Liouville's equation.

It is straightforward to check that ρ_{∞} is a steady state solution of (21), however, it remains to show that

$$\lim_{t \to \infty} \rho(x, t) = \rho_{\infty}(x), \tag{23}$$

irrespective of how the system is prepared. Moreover, ρ_{∞} is the unique steady solution of (21). To show this, we rewrite the Fokker-Planck equation in operator form

$$\frac{\partial \rho}{\partial t} = \mathcal{L}^* \rho, \tag{24}$$

again with reflecting boundary conditions

$$\mathcal{D}(\mathcal{L}^*) = \left\{ \rho \in H^2(\mathcal{X}) \mid \left(\gamma^{-1} \nabla \rho + \rho \nabla V \right) \cdot n = 0 \text{ on } \partial \mathcal{X} \right\}.$$
(25)

It can be shown that the eigenvalues of $-\mathcal{L}^*$ are real, nonnegative and can be ordered $0 = \lambda_0 < \lambda_1 < \lambda_2 < \cdots$. The corresponding eigenfunctions $\{\psi_n\}_{n=0}^{\infty}$ form a complete orthonormal basis of $\mathcal{D}(\mathcal{L}^*)$, and the zero eigenvalue is associated with the steady state solution $\psi_0 = \rho_{\infty}$. The reader is referred to Lemma 3.1 in [19] for a proof of these results. Now if we consider (24) for an initial condition $\rho_0(x) = c\psi(x)$, where ψ is an eigenfunction of $-\mathcal{L}^*$ and $c \in \mathbb{R}$, we can deduce that

$$\rho(x,t) = ce^{-\lambda t}\psi(x).$$
(26)

Therefore, by expanding an initial condition in terms of its eigenfunctions

$$\rho_0(x) = \sum_{n=0}^{\infty} c_n \psi(x), \qquad c_n = \int_{\mathcal{X}} \psi_n(x) \rho_0(x) dx.$$
(27)

we obtain

$$\rho(x,t) = \sum_{n=0}^{\infty} c_n e^{-\lambda_n t} \psi_n(x).$$
(28)

Then convergence to equilibrium as described by (23) can be seen by noting that $c_0 > 0$, and recalling that $\lambda_0 = 0$ is the only non-positive eigenvector. Lastly, the free energy converges to the unique minimum:

$$\lim_{t \to \infty} F(\rho(\cdot, t)) = F(\rho_{\infty}), \tag{29}$$

thereby completing the thermodynamic analogy.

4. The London Retail System

We demonstrate our model on the London retail system [11,20–22], with N = 625 residential wards and M = 49 shopping centres. We only include Metropolitan or International town centres in our analysis, as the smaller town centres carry a different type of retail activity and are better

modelled separately. For computationally efficient model calibration we let $\gamma \to \infty$ and assume that the process is confined to a domain \mathcal{X} containing a single minima, denoted x^* , of V. This approach is particularly efficient, since as $\gamma \to \infty$, ρ_{∞} collapses to a Dirac mass located at x^* , and we avoid the need to estimate the normalizing constant Z. Moreover, x^* can be found with a single run of a gradient-based optimization algorithm.

In this setting, we can find a suitable value for δ by noting that a zone with no inward flows satisfies

$$\delta = \kappa W_{\min},\tag{30}$$

where W_{\min} is the smallest size the system will support. Then by assuming all sizes sum to *K*, at equilibrium, we can obtain the following expression for κ

$$\kappa = K^{-1} \left[\delta M + \sum_{i=1}^{N} O_i \right]. \tag{31}$$

Lastly, we calibrate α and β values from the observation data comprising of log-transformed sizes $y = (y_1, \dots, y_M)^T$. The observation data is illustrated in Appendix (B). We assume that the logarithm of each size is given by a realization of ρ_{∞} , plus independent and identically centred Gaussian observation noise. The maximum likelihood values for α and β can then be obtained by solving the least squares optimization problem

$$\hat{\alpha}, \hat{\beta} = \underset{\alpha,\beta}{\operatorname{argmin}} |y - x_{a,\beta}^*|^2.$$
(32)

We solve (32) via a 2*D* grid search to obtain $\hat{\alpha} = 1.14$ and $\hat{\beta} = 0.54$. We provide further output of the calibration procedure in Appendix B.

We initialize the system at the calibrated value, so that $x_0 = x_{\hat{k},\hat{\beta}}^*$ and $\rho(x,0) = \delta(x-x_0)$. We then investigate two noise regimes with $\sigma = 0.02$ and $\sigma = 0.04$. For each regime, we simulate 1 000 trajectories by approximating (17) using the tamed Euler method [23,24]. We assume that the simulations take place on a larger domain where reflection does not occur within the finite simulation times. We compute Monte Carlo expectations to estimate the 5th and 95th percentiles and present the results in Figure. 1. We see that for $\sigma = 0.02$, the urban system does not escape the initial well within the simulation time. The percentiles at t = 10 are indicative of the equilibrium distribution providing there is a reflective boundary surrounding the well. For $\sigma = 0.04$, the additional noise provides enough energy for the system sometimes escape the initial well, and there is much greater uncertainty at t = 10. In this case the simulation is sensitive to the choice of domain. For both cases, the increase in uncertainty over time is a result of the second law of thermodynamics, which says that uncertainty must increase over time except for when there is a spontaneous decrease in potential energy.



Figure 1. Forward evolution of the London retail system for two noise regimes. Each column shows a different noise regime, and each row shows a different time. The inner and outer red rings show 5% and 95% percentiles, respectively. The light blue circles show the residential areas.

5. Further Work

In this article we have established a connection between the Fokker-Planck equation of a stochastic reformulation of the Harris and Wilson model and entropy maximization. The connection is made by way of a thermodynamic analogy. Whilst we have described a stochastic process that satisfies a second law of thermodynamics and has the desired equilibrium distribution, there are infinitely many other processes that we could have chosen from. A much broader question is the inverse problem: for a given equilibrium distribution, which process best describes the dynamics? So far we have considered the case that the first-order moment is constrained at equilibrium, however, enforcing the constraint for all times may lead to an improved model. The constrained Fokker-Planck equation is non-linear and non-local, and can allow for phase transitions [25,26]. In our future work we will investigate the constrained Fokker-Planck equations in the context of urban modelling.

Supplementary Materials: The code and data used for the case study in this manuscript can be found at the following repository: https://github.com/lellam/cities_and_regions.

Acknowledgments: L.E. was supported by EPSRC [EP/P020720/1]. MG was supported by EPSRC [EP/J016934/3, EP/K034154/1, EP/P020720/1, EP/R018413/1], an EPSRC Established Career Fellowship and The Alan Turing

Institute – Lloyd's Register Foundation Programme on Data-Centric Engineering. G.P. was supported by EPSRC [EP/P031587, EP/L020564, EP/L024926, EP/L025159]. A.W. was supported by EPSRC [EP/M023583/1]. This work was supported by The Alan Turing Institute under the EPSRC grant EP/N510129/1.

Author Contributions: L.E., G.P., M.G., A.W. designed the research, performed the research and prepared the article. L.E. performed the numerical experiments.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. Jaynes, E.T. Information theory and statistical mechanics. *Physical review* **1957**, *106*, 620.
- 2. Wainwright, M.J.; Jordan, M.I.; others. Graphical models, exponential families, and variational inference. *Foundations and Trends (in Machine Learning* **2008**, *1*, 1–305.
- 3. Wilson, A.G. A statistical theory of spatial distribution models. *Transportation research* **1967**, *1*, 253–269.
- 4. Wilson, A. Entropy in Urban and Regional Modelling: Retrospect and Prospect. *Geographical Analysis* **2010**, 42, 364–394.
- 5. Wilson, A.G. A family of spatial interaction models, and associated developments. *Environment and Planning A* **1971**, *3*, 1–32.
- 6. Harris, B.; Wilson, A.G. Equilibrium values and dynamics of attractiveness terms in production-constrained spatial-interaction models. *Environment and planning A* **1978**, *10*, 371–388.
- Dearden, J.; Wilson, A. A Framework for Exploring Urban Retail Discontinuities. *Geographical Analysis* 2011, 43, 172–187.
- 8. Davies, T.P.; Fry, H.M.; Wilson, A.G.; Bishop, S.R. A mathematical model of the London riots and their policing. *Scientific reports* **2013**, *3*, 1303.
- 9. Wilson, A. The "thermodynamics" of the city. In *Complexity and Spatial Networks*; Springer, 2009; pp. 11–31.
- 10. Wilson, A. Boltzmann, Lotka and Volterra and spatial structural evolution: an integrated methodology for some dynamical systems. *Journal of The Royal Society Interface* **2008**, *5*, 865–871.
- 11. Ellam, L.; Girolami, M.; Pavliotis, G.; Wilson, A. Stochastic Modelling of Urban Structure **2018**.
- 12. Mackey, M.C. *Time's arrow: the origins of thermodynamic behavior;* Courier Corporation, 2003.
- 13. Dougherty, J. Explaining statistical mechanics. *Studies in History and Philosophy of Science Part A* **1993**, 24, 843–866.
- 14. Dougherty, J.P. Foundations of non-equilibrium statistical mechanics. *Phil. Trans. R. Soc. Lond. A* **1994**, 346, 259–305.
- 15. Wilson, A.; Dearden, J. *Explorations in urban and regional dynamics: a case study in complexity science;* Routledge, 2015.
- 16. Wilson, A.G. *Entropy in urban and regional modelling*; Vol. 1, Routledge, 2011.
- 17. Vorst, A.C.F. A stochastic version of the urban retail model. *Environment and Planning A* 1985, 17, 1569–1580.
- 18. Lions, P.L.; Sznitman, A.S. Stochastic differential equations with reflecting boundary conditions. *Communications on Pure and Applied Mathematics* **1984**, *37*, 511–537.
- 19. Breiten, T.; Kunisch, K.; Pfeiffer, L. Control strategies for the Fokker-Planck equation. *ESAIM: Control, Optimisation and Calculus of Variations* **2017**.
- 20. Authority, G.L. Ward Profiles and Atlas. data.london.gov.uk, 2016.
- 21. Authority, G.L. Household Income Estimates for Small Areas. data.london.gov.uk, 2015.
- 22. Authority, G.L. Statistical GIS Boundary Files for London. data.london.gov.uk, 2016.
- 23. Hutzenthaler, M.; Jentzen, A.; Kloeden, P.E.; others. Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *The Annals of Applied Probability* **2012**, 22, 1611–1641.
- 24. Wang, X.; Gan, S. The tamed Milstein method for commutative stochastic differential equations with non-globally Lipschitz continuous coefficients. *Journal of Difference Equations and Applications* **2013**, *19*, 466–490.
- 25. Eberle, S.; Niethammer, B.; Schlichting, A. Gradient flow formulation and longtime behaviour of a constrained Fokker–Planck equation. *Nonlinear Analysis* **2017**, *158*, 142–167.
- 26. Mohammadi, M.; Borzi, A. A Hermite spectral method for a Fokker-Planck optimal control problem in an unbounded domain. *International Journal for Uncertainty Quantification* **2015**, *5*.

Appendix Decreasing Free Energy Functional

The free energy functional can be rewritten in terms of a function *h* that satisfies $\rho = h\rho_{\infty}$. By substituting the definition of ρ into the Fokker-Planck equation, we obtain the backward Kolmogorov equation for *h*

$$\frac{\partial h}{\partial t} = -\nabla V \cdot \nabla h + \gamma^{-1} \Delta h, \qquad x \in \mathcal{X}$$
(A1)

with the pure Neumann boundary condition

$$\nabla h \cdot n = 0, \qquad x \in \delta \mathcal{X},$$
 (A2)

and the initial condition $h(x,0) = \rho_0(x)\rho_{\infty}^{-1}(x)$. Then starting with the definition of F(h), we can substitute the time derivative for the backward Kolmogorov equation, integrate by parts and make use of the no flux boundary condition to see the claimed result:

$$\begin{split} \frac{d}{dt}F(h) &= \gamma^{-1}\frac{d}{dt}\int (h\ln h - h + 1)\rho_{\infty}dx, \\ &= \gamma^{-1}\int \ln h\frac{\partial h}{\partial t}\rho_{\infty}dx, \\ &= \gamma^{-1}\int \ln h \left[-\nabla V\cdot\nabla h + \gamma^{-1}\Delta h\right]\rho_{\infty}dx, \\ &= \gamma^{-1}\int \ln h \left(-\nabla V\cdot\nabla h\right)\rho_{\infty}dx + \gamma^{-1}\int \ln h(\Delta h)\rho_{\infty}dx, \\ &= \gamma^{-1}\int \ln h \left(-\nabla V\cdot\nabla h\right)\rho_{\infty}dx + \gamma^{-1}\int \ln h(\nabla h\cdot n)dx - \gamma^{-1}\int \nabla(\rho_{\infty}\ln h)\cdot\nabla hdx, \end{split}$$
(A3)
$$&= -\gamma^{-1}\int (\nabla\ln h\cdot\nabla h)\rho_{\infty}dx, \\ &= -\gamma^{-1}\int h^{-1}|\nabla h|^{2}\rho_{\infty}dx, \\ &= -\int \rho \left|\nabla\frac{\delta F}{\delta\rho}\right|^{2}dx, \\ &\leq 0. \end{split}$$

Appendix Model Calibration

We determine α and β values by solving the maximum likelihood problem in (32) using a 1,000 × 1,000 grid search. We use an appropriate scaling of the cost matrix so that it is reasonable to perform the grid search over $[0,2]^2$. We specify K = 1 and normalize the observation data so that the sizes sum to 1. We then specify $\kappa = 1.44$ and $\delta = 0.009$ in accordance with (30)- (31). The results of the grid search are shown in Figure. A1, for which it can be seen that the likelihood is maximized in the region $1 < \alpha < 2$. The fitted values are $\hat{\alpha} = 1.53$ and $\hat{\beta} = 0.66$. The discontinuities in the parameter space are in line with those discussed in [7]. We plot the value of *y* against $x^*_{\hat{\alpha},\hat{\beta}}$ in Figure. A2 from which we conclude that the model gives a reasonable fit.



Figure A1. Evaluations of the log-likelihood, given by the negative of the objective function in (32), over a grid of $1,000 \times 1,000$ values of α and β .



Figure A2. The observed values *y*, comprising of log-transformed sizes, against the predicted value $x^*_{\alpha,\beta}$.

© 2018 by the authors. Submitted to *Proceedings* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).