



Article

Electrical Power Diversification: An Approach Based on the Method of Maximum Entropy in the Mean

Rafael Bautista ¹, Henryk Gzyl ², Enrique ter Horst ^{1,*} and Germán Molina ³¹ School of Management, Universidad de los Andes, Bogotá 111711, Colombia; rbautist@uniandes.edu.co² Centro de Finanzas, IESA, Caracas 1011, Venezuela; henryk.gzyl@iesa.edu.ve³ Idalion Capital Group, Quantitative Trading, 12 Hay Hill, London W1J 8NR, UK; german@germanmolina.com

* Correspondence: ea.terhorst@uniandes.edu.co

Abstract: Electrical energy is generated in different ways, each located at some specific geographical area, and with different impact on the environment. Different sectors require heterogeneous rates of energy delivery, due to economic requirements. An important problem to solve is to determine how much energy must be sent from each supplier to satisfy each demand. Besides, the energy distribution process may have to satisfy ecological, technological, or economic cost constraints.

Keywords: electricity supply and demand; energy supply diversification; contamination constraints; inverse problems; maximum entropy in the mean



Citation: Bautista, R.; Gzyl, H.; ter Horst, E.; Molina, G. Electrical Power Diversification: An Approach Based on the Method of Maximum Entropy in the Mean. *Entropy* **2021**, *23*, 281. <https://doi.org/10.3390/e23030281>

Received: 21 November 2020

Accepted: 14 December 2020

Published: 26 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The problem of optimal and well-diversified energy distribution in a given country is central to the well-functioning of economic and social infrastructures. There are two aspects to the problem. On the one hand, we have to consider the problem of satisfying the demand without unnecessary surplus (unless there is a market for the surplus). On the other hand, we must ensure that the resulting distribution is well-diversified.

The first problem is a typical exercise of demand-supply with given cost constraints. Costs include economic, ecological, and social costs. Any cost that can be translated into a constraint on the energy flow between consumer and supplier (or source and sink) should be taken into account. This problem has received considerable attention in the applied mathematics literature when the data about the amounts produced are exact. A review of the literature is available in [1].

In addition, as there might be uncertainties in the demand and/or the supply, it is realistic to suppose that the data are specified up to intervals. Of course, there are obvious interval constraints of this type: a given supplier cannot produce more than a certain amount at a given site. However, the amount produced may fluctuate within a range within the maximum capacity. Similar limitations can be observed on the demand side.

Linear programming techniques are one of the standard tools to deal with the problem, especially when there is a monetary cost that must be minimized. The main goal of this paper is to propose the method of maximum entropy on the mean to deal with the transportation problem with cost constraints and range constraints on the solutions. The method that we propose is based on an extension of the standard method of maximum entropy proposed in [2]. Attempts to solve the problem using maximum entropy based methods are long-dated (see Chapter 13 in Kapur's [3] and the references in that chapter for earlier work, in particular to the work by Wilson [4]). An interesting application can be found in [5]. Our approach to the use of the method of maximum entropy is completely different.

The problem of diversification is reviewed in [6,7], where more references are listed, and we bring in the entropy-based measure proposed there to measure the diversification of the solution to the problem considered above.

To state the problems treated here, consider M sources of electricity and denote by $\{S_1, \dots, S_M\}$ the amount supplied by each source. The labeling addresses the possibility that in the same geographical area there may exist two different types of electricity sources, which would be labeled by two different subscripts.

Consider N different electricity consumers, each of whom requiring $\{D_1, \dots, D_N\}$ units of energy per unit time. These may describe geographical areas and/or different types of consumer within a given area. For example, within a geographical area, there may exist large chemical industries or mineral processing plants that may need large amounts of electricity, besides what is needed for domestic use by the local population.

The first group of constraints is imposed by the nature of the generation process or by the local demand of electricity: We suppose that the demand D_i at each site $i = 1, \dots, N$ and the supply S_j at site $j = 1, \dots, M$ are both known. Demand at site i represents the demand by a type of energy consumer which might be distributed geographically, but which shares some type of collective characteristic, such as being a domestic (household) consumer in a geographic area.

1.1. Problem Statements

If there were no other constraints, the simplest problem of distributing the electricity could be stated as follows.

Problem 1 (First electricity supply-demand problem). Denote by $\{x_{i,j} : i = 1, \dots, N; j = 1, \dots, M\}$ the amount of electricity required at site i and produced at site j per unit time. We suppose that the network is such that the every electricity-producing site is connected to every consuming site. Otherwise, we set the corresponding $x_{i,j} = 0$ without further ado.

The electricity matching problem consists of determining values $\{x_{i,j} : i = 1, \dots, N; j = 1, \dots, M\}$ that satisfy the following demand-supply constraints:

$$\sum_{j=1}^M x_{i,j} = D_i, \quad \text{for } i = 1, \dots, N \quad \text{and} \quad \sum_{i=1}^N x_{i,j} = S_j \quad \text{for } j = 1, \dots, M. \quad (1)$$

Comments:

- (i) Production per unit time refers to an average produced or required during some standardized time interval (one hour, for example).
- (ii) These constraints can be replaced by intervals to allow for uncertainty in the demand or uncertainty in the supply. We describe this further below.
- (iii) Note that we are taking into account possible nonlinear constraints, resulting from the actual physical transport of energy through the network, in which the losses could depend on the amount of energy being transported.

In practice, cost constraints may also exist. Here, we suppose that costs are regulated and fixed (by competition or government agencies) and fairly passed on to consumers, but these constraints have to be taken into account.

There might also exist regulatory environmental constraints. Each mode of electricity generation has an environmental impact, measured by C_i grams of contaminant per unit of electricity generated by the i th supplier per unit time. Therefore, the total amount of contaminants generated by the supplier is

$$\sum_{i=1, j=1}^{i=N, j=M} C_i x_{i,j} = \sum_{i=1}^N C_i \left(\sum_{j=1}^M x_{i,j} \right).$$

With this, instead of Problem 1, we consider now the following:

Problem 2 (Second electricity supply-demand problem). *With the notation introduced above, the electricity matching problem consists of determining $\{x_{i,j} : i = 1, \dots, N; j = 1, \dots, M\}$ that satisfy the demand-supply constraints (1) as well as cost constraints solving:*

$$\sum_{i=1, j=1}^{i=N, j=M} C_{r,(i,j)} x_{i,j} = K_r, \quad r = 1, \dots, R. \quad (2)$$

Here, K_r is a maximum cost (environmental impact) generated (or incurred) by the electricity-producing system that regulators allow. Most of the time the constraint will not depend on the connection (edge) (i, j) but only on the source (type of energy that is produced at the j th source). In this case, $C_{r,(i,j)}$ does not change as i changes.

For a combination of technological and economical reasons, we might be forced to consider flexible constraints. For example, to cover for possible downward or upward movement in demand or supply, we might consider those values to fall in a range. Similarly, we could include a tolerance for sudden fluctuations in the cost of production, or the demand at a certain node. Instead of point-valued constraints, we might replace both Problems 1 and 2 by:

Problem 3 (Third electricity supply-demand problem). *To allow for different types of constraints, we extend our previous notation. Denote by C a $R \times d$ -matrix, where $d = N \times M$, and denote its elements by $C_{r,(i,j)}$, where $1 \leq r \leq R$ labels the cost constraints (or cost restrictions). Now, instead of a point value K , we consider a range dataset given by:*

$$\mathcal{K} = \prod_{r=1}^R [K_1^r, K_2^r], \quad \text{with } K_1^r \leq K_2^r \text{ for } r = 1, \dots, R. \quad (3)$$

The problem to solve is to find $x_{i,j}$ in some given range such that:

$$Cx \in \mathcal{K}. \quad (4)$$

When \mathcal{K} reduces to a point, that is, when $\mathcal{K} = \{\mathbf{k}\}$ with $\mathbf{k} \in \mathbb{R}^R$, we have a problem with **point constraints**.

Problems 1–3 are ill-posed linear inverse problems subject to convex constraints. As such, each might have infinitely many possible solutions because the number of unknowns is usually significantly larger than that of the dataset. The convex constraints include positivity constraints as well as those in (3).

The method of maximum entropy in the mean is especially designed to deal with this type of problems. The representation of the solution is such that it allows for explicit sensitivity analysis. The standard maximum entropy method (SME) is the stepping stone towards the method of maximum entropy in the mean (MEM).

1.2. Contents of the Paper

Building on the aforesaid notation, we briefly describe the solution to Problems 1–3 in Section 2. The solution of the three problems, given by (21), looks the same, although what changes is the specification of the problem data. The actual derivation of the solutions of Problems 1 and 2 follows the same pattern, and Problem 3 uses the solution of Problem 2 as a stepping stone.

In Section 3, we recall the mathematical details of the procedure to arrive at the results listed in Section 2. We include a short digression on using the concept of entropy to quantify how diversified is a solution to the supply-demand problem. Further, we explicitly compute it for each of the examples.

In Section 4, we work out two illustrating toy examples which take into account all the essentials. In the first one, we consider only point data, while, in the second example,

we show the flexibility of the approach with some of the data in intervals. The natural production constraints are specified up to an interval to incorporate the possibility of fluctuation in the energy generation output.

2. Mathematical Model Derivation

Here, we collect the basics results about the SME method, MEM with point data, and MEM with data in ranges. We explain the resulting (21).

2.1. MEM for Point Data

The standard maximum entropy (SME) method was almost simultaneously proposed by Jaynes [2] and Kullback [8]. It is a variational method to determine a probability density $\rho(\xi)$ such that $dP(\xi) = \rho(\xi)dQ(\xi)$ satisfying some integral constraints. MEM focuses on how to transform an ill-posed linear inverse problem with constraints upon the solutions into a problem of determining a probability density. This is explained in [9,10]. To continue, we need to introduce some notations.

Let $\Omega = \prod_{i=1}^d [a_i, b_i]$ and denote by \mathcal{F} the Borel subsets of Ω . Consider a (reference) measure $dQ(\xi)$ upon \mathcal{F} . Let $\mathcal{P}(Q) = \{P \text{ probability measure on } \Omega : dP(\xi) = \rho(\xi)dQ(\xi)\}$, that is the class of probabilities absolutely continuous with respect to Q . Let $X: \Omega \rightarrow \Omega$ denote the coordinate or identity mapping, that is $X(\xi) = \xi$. Let C be a $K \times d$ -matrix and let $y \in \mathbb{R}^K$ be some given vector. The constraint set is defined by:

$$\mathcal{C}(Q, y) = \{P \in \mathcal{P} \mid CE_P[X] = y\}. \quad (5)$$

This is a convex set. The best way to choose a point in it is to maximize some meaningful concave function defined on it.

Definition 1. The entropy of $P \in \mathcal{P}$ is defined by

$$S_Q(P) = - \int_{\Omega} \rho(\xi) \ln \rho(\xi) dQ(\xi). \quad (6)$$

whenever $\int_{\Omega} \rho(\xi) |\ln \rho(\xi)| dQ(\xi) < \infty$, and equal to $-\infty$ otherwise.

Then, we can state the entropy maximization problem:

Problem 4. Find $\rho^*(\xi)$ at which $S_Q(P)$ achieves its maximum subject to the constraints (5). That is, determine

$$\rho^* = \operatorname{argsup}\{S_Q(P) \mid P \in \mathcal{C}(Q, y)\}.$$

Once $dP^* = \rho^* dQ$ has been determined, the solution to the algebraic problem is given by

$$x^* = E_{P^*}[X] = \int_{\Omega} \xi \rho^*(\xi) dQ(\xi) \quad (7)$$

which clearly satisfies $x^* \in \Omega$ and $Cx^* = y$. In our present case, a direct application of the method of Lagrange multipliers leads to the representation

$$\rho^*(\xi) = \frac{e^{-(C^t \lambda^*, \xi)}}{Z(\lambda^*)} \quad (8)$$

where the normalization factor is given by

$$Z(\lambda) = \int_{\Omega} e^{(C^t \lambda, \xi)} dQ(\xi), \quad \lambda \in \mathbb{R}^K. \quad (9)$$

The optimal Lagrange multiplier λ^* is to be found minimizing the dual entropy function

$$\Sigma(\lambda, \mathbf{y}) = \ln Z(\lambda) + \langle \lambda, \mathbf{y} \rangle \tag{10}$$

over $\lambda \in \mathbb{R}^K$. We use $\langle \lambda, \mathbf{y} \rangle$ to denote the standard scalar. The connection between the primal and the dual problems is explained in [11,12]. The importance of this connection is that it transforms the search over an infinite-dimensional space (the class of densities) onto the problem of finding a minimum of a convex function over a finite-dimensional space.

Below, we make use of the fact that the entropy of ρ^* is given by:

$$S_Q(P^*, \mathbf{y}) = \ln Z(\lambda^*) + \langle \lambda^*, \mathbf{y} \rangle = \inf\{\ln Z(\lambda) + \langle \lambda, \mathbf{y} \rangle\} \tag{11}$$

where the label specifying the point constraint is added for emphasis and used below.

2.2. MEM for Data in Ranges

The extension of the procedure described in the previous section to Problem (20) goes as follows. Instead of (5), we now have

$$\mathcal{C}(Q, \mathcal{K}) = \{P \in \mathcal{P} \mid CE_P[\mathbf{X}] \in \mathcal{K}\} = \bigcup_{\mathbf{y} \in \mathcal{K}} \mathcal{C}(Q, \mathbf{y}). \tag{12}$$

Therefore, we consider the extended entropy maximization problem:

Problem 5. Find $\rho^*(\xi)$ at which $S_Q(P)$ achieves its maximum subject to the constraints (12). That is, determine

$$\rho^* = \operatorname{argsup}\{S_Q(P) \mid P \in \mathcal{C}(Q, \mathcal{K})\}.$$

Considering (12), we can rewrite (5) as

$$\rho^* = \operatorname{argsup}\{\sup\{S_Q(P) \mid P \in \mathcal{C}(Q, \mathbf{y})\}, \mid \mathbf{y} \in \mathcal{K}\} = \operatorname{argsup}\{S_Q(P^*, \mathbf{y}) \mid \mathbf{y} \in \mathcal{K}\}$$

Now, making use of (11), we rewrite the last identity as

$$\sup\left\{\inf\left\{\ln Z(\lambda) + \langle \lambda, \mathbf{y} \rangle \mid \lambda \in \mathbb{R}^K\right\} \mid \mathbf{y} \in \mathcal{K}\right\}.$$

We can exchange the maximization with the minimization to obtain

$$\inf\left\{\ln Z(\lambda) + \sup\{\langle \lambda, \mathbf{y} \rangle \mid \mathbf{y} \in \mathcal{K}\} \mid \lambda \in \mathbb{R}^K\right\}.$$

To compute the inner maximization, we perform the following affine transformation

$$y_r = \frac{K_2^r - K_1^r}{2} \zeta_r + \frac{K_2^r + K_1^r}{2} = \frac{K_2^r - K_1^r}{2} \left[\zeta_r + \frac{K_2^r + K_1^r}{K_2^r - K_1^r} \right] \quad r = 1, \dots, R$$

where $\zeta_r \in [-1, 1], r = 1, \dots, R$. For any $\lambda \in \mathbb{R}^K$ we have

$$\sup\{\lambda_r, \zeta_r \mid \zeta_r \in [-1, 1]\} = |\lambda_r|.$$

Therefore, we obtain that ρ^* is as (8), that is,

$$\rho^*(\xi) = \frac{e^{-\langle \mathbf{c}^t \lambda^*, \xi \rangle}}{Z(\lambda^*)} \tag{13}$$

but this time λ^* is to be found as the minimizer of

$$\Sigma(\lambda, \mathcal{K}) = \ln Z(\lambda) + \sum_{r=1}^R \left(\frac{K_2^r - K_1^r}{2} \right) |\lambda_r| + \sum_{r=1}^R \left(\frac{K_2^r + K_1^r}{2} \right) \lambda_r. \tag{14}$$

2.3. Computation of $Z(\lambda)$

The last step in the mathematical procedure consists of the specification of the normalization factor $Z(\lambda)$. In the current setup, in which the constraint space is a Cartesian product of intervals, that is, $\Omega = \prod_{i=1}^d [a_i, b_i]$, the choice of reference measure is very simple: just set

$$dQ(\xi) = \prod_{j=1}^d \left(\epsilon_{a_j}(d\xi_j) + \epsilon_{b_j}(d\xi_j) \right)$$

where we use $\epsilon_a(d\xi)$ for the point mass at a (sometimes written $\delta(\xi - a)d\xi$). The intuition is that any point in $[a_i, b_i]$ is the expected value with respect to a probability carried by the end points of the interval. With this, the function (9) is given by

$$Z(\lambda) = \int_{\Omega} e^{-\langle C^t \lambda, \xi \rangle} dQ(\xi) = \prod_{j=1}^d \left(e^{-(C^t \lambda)_j a_j} + e^{-(C^t \lambda)_j b_j} \right). \tag{15}$$

2.3.1. The Case of Point Constraints

When (4) consists of determining the energy supply pattern for a given point constraint \mathbf{k} , the dual entropy (10) now becomes

$$\Sigma(\lambda, \mathbf{k}) = \sum_{j=1}^d \ln \left(e^{-(C^t \lambda)_j a_j} + e^{-(C^t \lambda)_j b_j} \right) + \langle \lambda, \mathbf{k} \rangle. \tag{16}$$

This function is clearly defined, strictly convex in λ , and, since it tends to ∞ as $\|\lambda\| \rightarrow \infty$, it has a unique minimizer. In addition, from the first-order condition for λ^* to be a minimizer, it is easy to see that

$$x_j^* = a_j \frac{e^{-(C^t \lambda^*)_j a_j}}{e^{-(C^t \lambda^*)_j a_j} + e^{-(C^t \lambda^*)_j b_j}} + b_j \frac{e^{-(C^t \lambda^*)_j b_j}}{e^{-(C^t \lambda^*)_j a_j} + e^{-(C^t \lambda^*)_j b_j}}. \tag{17}$$

The weights of a_j and b_j are positive and add up to 1 and can therefore be interpreted as the (maxentropic) probability that the auxiliary random variable X assumes those values, and \mathbf{x}^* is its expected value. They satisfy the constraint $C\mathbf{x}^* = \mathbf{k}$, as evidenced from the first order condition for λ^* to be a minimizer.

2.3.2. The Case of Data in Ranges

When the data are specified up to a range as in the previous section, the aforescribed approach can also be used. The function $Z(\lambda)$ is as above, but this time the dual entropy is different. According to the methods in Section 2 the dual entropy this time is

$$\Sigma(\lambda, \mathcal{K}) = \sum_{j=1}^d \ln \left(e^{-(C^t \lambda)_j a_j} + e^{-(C^t \lambda)_j b_j} \right) + \sum_{r=1}^R \left(\frac{K_2^r - K_1^r}{2} \right) |\lambda_r| + \sum_{r=1}^R \left(\frac{K_2^r + K_1^r}{2} \right) \lambda_r. \tag{18}$$

Again, this is a strictly convex function of λ that tends to ∞ as $\|\lambda\| \rightarrow \infty$. Thus, it has a unique minimizer at some point λ^* . Once this point has been found, the solution can be represented by (17), and now we have that $C\mathbf{x}^* \in \mathcal{K}$.

3. The Maxentropic Solution to the Energy Diversification Problem

In this section, we see how to relabel the unknowns, describe the maxentropic solution to the resulting problem and describe how to compute the energy diversification index from the resulting solution.

3.1. Statement of the Problems and Representation of the Solution

Relabeling the Problem

Here, we relabel the unknowns and the data, not only to simplify the presentation, but also to allow the numerical implementation. We apply the following anti-lexicographic labeling $\ell : \{1, 2, \dots, N\} \times \{1, 2, \dots, M\} \rightarrow \{1, 2, \dots, d = NM\}$ specified as follows:

$$\begin{aligned} (i, 1) &\rightarrow i, & \text{for } i = 1, \dots, N, \text{ \& } j = 1 \\ (i, 2) &\rightarrow i + N, & \text{for } i = 1, \dots, N, \text{ \& } j = 2 \\ & & \vdots \\ (i, M) &\rightarrow i + (M - 1)N, & \text{for } i = 1, \dots, N, \text{ \& } j = M \end{aligned}$$

Similarly, let us write $x_i \in [a_i, b_i]$ to denote the constraints upon the unknowns in the new labeling. Notice that each constraint might be repeated a number of times since it is an energy production constraint and each source appears repeated as many times as its connections to demand centers.

After the relabeling of the edges of the supply-demand system, the constraint matrix C can be relabeled in the obvious way $C_{r,(i,j)} \rightarrow C_{r,\ell(i,j)}$ and our problem becomes

Problem 6 (Generic problem). Find $x \in \prod_{j=1}^d [a_j, b_j]$ such that either

$$Cx = k \tag{19}$$

when we have point data or

$$Cx \in \mathcal{K}. \tag{20}$$

when the data are specified up to a range.

In general, supply-demand problems lead to singular constraint matrices C . This is essentially due the fact that $\sum D_i = \sum S_j = T$, which amounts to saying that there are two sets of rows in C which have the same sum, therefore they are linearly related or C is singular. As examples consider, the two matrices in Section 3.

The solution to Problems (19) and (20) is shown to be given by (17), which we display as

$$x_j^* = a_j \frac{e^{-(C^t \lambda^*)_j a_j}}{e^{-(C^t \lambda^*)_j a_j} + e^{-(C^t \lambda^*)_j b_j}} + b_j \frac{e^{-(C^t \lambda^*)_j b_j}}{e^{-(C^t \lambda^*)_j a_j} + e^{-(C^t \lambda^*)_j b_j}}. \tag{21}$$

The difference between its applicability to (19) and (20) comes from the fact that in each case the λ^* appearing in (21) is found minimizing a different (dual entropy) convex function. Observe that x_j^* is a convex combination of weights adding up to 1, and therefore $x_j^* \in [a_j, b_j]$ for $j = 1, \dots, d$. Note that either $Cx^* = k$ in the case of point constraints or $Cx^* \in \mathcal{K}$ in the case of extended constraints.

3.2. The Energy Diversification According to Stirling's Measure

To analyze how well-diversified our numerical solutions turn out to be, we invoke Stirling's proposal of computing the entropy of a quantity directly related to the solution, namely the energy transferred to site i from source j per unit of total demand. For this, note

that the total demand $T = \sum_{i=1}^N D_i$ can also be written as $T = \sum_{\{i,j\}} x_{i,j}$. Define now the energy transferred per unit demand by

$$p_{i,i} = \frac{x_{i,j}}{T} = \frac{x_{i,j}}{\sum_{\{i,j\}} x_{i,j}} \quad (i, j) \in \{1, \dots, N\} \times \{1, \dots, M\}. \tag{22}$$

Conversely, using the lexicographic labeling,

$$p_k = \frac{x_k}{T} = \frac{x_k}{\sum_n x_n} \quad k \in \{1, \dots, NM\}. \tag{23}$$

Observe that with this rescaling, p_k are a true probability distribution.

The Stirling energy diversification coefficient is defined as the Boltzmann-Gibbs-Shannon entropy of p , namely:

$$ED(x) = - \sum_k^{NM} p_k \ln p_k. \tag{24}$$

To check how well-diversified the maxentropic solution x^* given by (21) is, we can compare $ED(x^*)$ versus the most diversified probability distribution on $\{1, \dots, NM\}$, namely the uniform distribution which has an entropy $\ln NM$.

4. Numerical Examples

Here, we implement the results obtained in the previous section in two illustrating toy examples. In the first one, we consider the simplest case of a global aggregate demand, and, in the second, we consider two possible demands: domestic and industrial.

4.1. Case 1: Aggregated Demand

The sources of energy, the bounds on their outputs, and their technological cost constraints are given in Table 1. All of these costs are related to the emission of combustion gases. It might be interesting to find ways of quantifying a technological (environmental) cost for the electricity produced in nuclear plants. We mention as well that one could add solar, eolic, sea wave, or geothermal energy. None of them contributes to the greenhouse effect.

Table 1. Toy dataset.

Type	Source	Lower Bd. (a)	Upper Bd. (b)	Cost
1	Hydraulic	0.01	0.8	0.00
2	Gas	0.01	0.5	0.37
3	Fuel	0.01	0.7	0.76
4	Coal	0.01	0.4	0.96

We consider a uniform nonzero lower bound in all cases to accommodate the possibility that no source fails completely. Otherwise, we can set the corresponding a_j equal to 0. The upper bounds can be thought to be given as fraction of a total possible maximal demand.

We consider two constraints, one a total demand constraint and the other an ecological or regulatory cost constraint. Thus, the constraint matrix becomes:

$$C = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0.0 & 0.3 & 0.7 & 0.9 \end{bmatrix}.$$

This matrix is clearly singular. The vector of unknowns $\mathbf{x}^t = (x_1, x_2, x_3, x_4)$ is constrained to be in $\Omega = [0.01, 0.8] \times [0.01, 0.5] \times [0.01, 0.7] \times [0.01, 0.4]$. As point data vector, we consider $\mathbf{k}^t = (1.8, 0.5)$. Thus, the problem that we have to solve is

$$\text{Find } \mathbf{x} \in \Omega \text{ such that } \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0.0 & 0.3 & 0.7 & 0.9 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1.2 \\ 0.5 \end{bmatrix}.$$

Once the optimal λ^* is available, the quantities to be supplied by the sources are computed as shown in (21). The optimal solution is:

$$\mathbf{x}^* = (0.459, 0.253, 0.305, 0.181)$$

The diversification measure

With the entropic solution to Problem (19), we form $p_j = x_j^*/k_1$, which is the proportion of the available energy supply to be provided by the j th supplier, and, with that, we proceed as described in Section 2.2 and compute the entropic diversification measure as:

$$ED(\mathbf{x}^*) = - \sum_{j=1}^4 p_j \ln p_j. \quad (25)$$

as proposed by Stirling, that is:

$$ED(\mathbf{x}^*) = - \sum_{j=1}^4 p_j \ln p_j = 1.329.$$

This should be compared with the entropy of the uniform distribution, which in this case is $\ln 4 = 1.386$. The relative error is about 4%.

4.2. Case 2: Disaggregated Demand

To continue with the toy examples, we now suppose that electrical power has to be supplied to two types of consumers: domestic and industrial. Therefore, the supply of the i th source is split into two components x_i, x_{4+i} , which means that our vector of unknowns is an eight-dimensional vector $\mathbf{x}^t = (x_1, x_2, \dots, x_4, \dots, x_8)$. The first four components describe how much of the output of the i th source goes into domestic demand, and the last four how much goes into industrial demand. This implies that the supply of the i th source has to meet the additional demand constraint $x_i + x_{4+i} \leq b_i$, where the b_i s are given in Table 1. Let us suppose that we have one technological cost constraint and one production cost constraint. The matrix of constraints is now

$$\mathbf{C} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0.0 & 0.3 & 0.7 & 0.9 & 0.0 & 0.3 & 0.7 & 0.9 \\ 0.6 & 0.4 & 0.7 & 0.3 & 0.6 & 0.4 & 0.7 & 0.3 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Note that the third and fourth rows consist of two equal halves. The third contains the technological (environmental) cost constraint and the fourth contains the production cost constraint. Note as well that the sum of Rows 1 and 2 equals the sum of Rows 5–8, therefore the determinant of \mathbf{C} is 0. That is, even though the matrix \mathbf{C} is square, the problem is nevertheless ill-posed.

Let us now suppose that the aggregate domestic and industrial demands are, respectively, $k_1 = 0.9$; $k_2 = 1.05$. This time the total is intentionally higher than before. Let the technological constraint upon the energy production be $k_3 = 0.73$ while the production constraint is $k_4 = 1.06$. Now, instead of point data for the remaining constraints, we have $x_i + x_{4+i}$. Again, we are supposing that every supplier produces more than a minimum a_i and less than b_i . The range for the data in this example is the following:

$$\mathcal{K} = \{k_1\} \times \{k_2\} \times \{k_3\} \times \{k_4\} \times [K_1^5, K_2^5] \times [K_1^6, K_2^6] \times [K_1^7, K_2^7] \times [K_1^8, K_2^8].$$

For the numerical computations, we use:

$$\mathcal{K} = \{0.9\} \times \{1.05\} \times \{0.73\} \times \{1.06\} \times [0.65, 0.75] \times [0.35, 0.45] \times [0.65, 0.75] \times [0.25, 0.35].$$

Clearly, some of the ranges are degenerate (just a point) while the others are just the natural range for the corresponding supply. We still have to specify the range for the components of x . The obvious choice is $x_i, x_{4+i} \in [a_i, b_i]$, which suggests the following choice for the constraint (or sample) space Ω :

$$\Omega = [0.01, 0.8] \times [0.01, 0.7] \times [0.01, 0.5] \times [0.01, 0.4] \times [0.01, 0.8] \times [0.01, 0.7] \times [0.01, 0.5] \times [0.01, 0.4].$$

Note that we are assuming that the constraints upon the individual demands are the same as the constraint upon the output of each source. This facilitates the typography.

The problem to solve now becomes:

$$\text{Find } x \in \Omega \text{ such that } Cx \in \mathcal{K}.$$

As the previous case, the solution is also given by

$$x_j^* = a_j \frac{e^{-(C^t \lambda^*)_j a_j}}{e^{-(C^t \lambda^*)_j a_j} + e^{-(C^t \lambda^*)_j b_j}} + b_j \frac{e^{-(C^t \lambda^*)_j b_j}}{e^{-(C^t \lambda^*)_j a_j} + e^{-(C^t \lambda^*)_j b_j}} \quad j = 1, \dots, 8.$$

In this case, C is the 8×8 -matrix specified above. The vector λ^* is an eight-dimensional vector that is obtained minimizing the dual entropy (18)

$$\Sigma(\lambda, \mathcal{K}) = \sum_{j=1}^8 \ln(e^{-(C^t \lambda)_j a_j} + e^{-(C^t \lambda)_j b_j}) + \sum_{r=1}^8 \left(\frac{K_2^r - K_1^r}{2} \right) |\lambda_r| + \sum_{r=1}^8 \left(\frac{K_2^r + K_1^r}{2} \right) \lambda_r.$$

When we separate the degenerate from the non-degenerate intervals, the second and third summation become:

$$\sum_{i=1}^4 \lambda_i k_i + \sum_{r=5}^8 \left(\frac{K_2^r - K_1^r}{2} \right) |\lambda_r| + \sum_{r=5}^8 \left(\frac{K_2^r + K_1^r}{2} \right) \lambda_r.$$

Undoing the lexicographic ordering, the maxentropic solution in this case is

$x_{i,j}$	1	2	3	4
1	0.300	0.178	0.301	0.106
2	0.369	0.221	0.327	0.119

The diversification measure

Now that we have obtained the x_j^* solving our problem, we apply (23), form $p_j = x^* / (k_1 + k_2)$, and compute the entropic measure of diversity (24).

$$ED(x^*) = - \sum_{j=1}^8 p_j \ln p_j = - \sum_{i,j} \frac{x_{i,j}^*}{k_1 + k_2} \ln \frac{x_{i,j}^*}{k_1 + k_2} = 2.00.$$

Taking into account that the maximum diversification possible corresponds to a uniform probability distributions $q_j = 1/8$ and that $\ln 8 = 2.079$, the maxentropic solution is well diversified.

5. Conclusions and Policy Implications

The two illustrating toy examples considered show the viability of the method of maximum entropy in the mean to solve constrained energy transportation problems. To test the diversification of the transportation policies obtained, we rescaled the solution x^* to the energy transportation problem. That rescaling can be interpreted as the energy transported per unit of total demand. It renders the solution as a probability distribution on the edges of the transportation graph. The resulting entropy of the probability is a measure of diversification.

An interesting aspect of the method of maximum entropy in the mean is that policy requirements can be imposed as ab initio constraints on the solution.

In the two examples that we considered, the solution happened to be quite well diversified, but a general comparison is still lacking.

Author Contributions: All coauthors have worked equally in terms of conceptualization, methodology, software, validation, formal analysis, investigation, resources, data curation, writing-original draft preparation, and writing-reviewing and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

1. Ku, A. Modeling Uncertainty in Electricity Capacity Planning, Thesis, 1995. Available online: www.s3.amazonaws.com/academia.edu.documents/52911593/thesis_aku.pdf (accessed on 13 December 2020).
2. Jaynes, E. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630. [[CrossRef](#)]
3. Kapur, J.N. *Maximum Entropy Models in Science and Engineering*; Wiley: New York, NY, USA, 1998.
4. Wilson, A.G. *Entropy in Urban and Regional Modeling*; Pion: London, UK, 1970.
5. Xiao, X.Y.; Chao, M.A.; Yang, H.G.; Li, H.Q. Maximum entropy probability method applied to assess voltage sag frequency due to transmission line fault in the electric power system. *Appl. Stoch. Models Bus. Ind.* **2009**, *26*, 595–608. [[CrossRef](#)]
6. Stirling, A. Multicriteria diversity analysis. A novel heuristic framework for appraising energy portfolios. *Energy Policy* **2010**, *38*, 1622–1634. [[CrossRef](#)]
7. Stirling, A. Diversity and ignorance in electricity supply investment. *Energy Policy* **1994**, *22*, 195–216. [[CrossRef](#)]
8. Kullback, S. *Information Theory and Statistics*, 2nd ed.; Dover Publications: New York, NY, USA, 1955.
9. Golan, A.; Gzyl, H. A generalized maxentropic inversion procedure for noisy data. *Appl. Math. Comput.* **2002**, *127*, 249–260. [[CrossRef](#)]
10. Gzyl, H.; Velásquez, Y. *Linear Inverse Problems: The Maximum Entropy Connection*; World Scientific Publishers: Singapore, 2011.
11. Borwein, J.; Lewis, A. *Convex Analysis and Nonlinear Optimization*; CMS Books; Springer: New York, NY, USA, 2000.
12. Mead, L.R.; Papanicolau, N. Maximum entropy in the problem of moments. *J. Math. Phys.* **1984**, *25*, 2404–2417. [[CrossRef](#)]