



(12) **Offenlegungsschrift**

(21) Aktenzeichen: **10 2022 104 313.4**

(22) Anmeldetag: **23.02.2022**

(43) Offenlegungstag: **24.08.2023**

(51) Int Cl.: **G06N 3/08 (2023.01)**

**B60L 15/20 (2006.01)**

(71) Anmelder:  
**Dr. Ing. h.c. F. Porsche Aktiengesellschaft, 70435  
Stuttgart, DE**

(56) Ermittelter Stand der Technik:  
**DE 10 2020 118 805 A1**

(72) Erfinder:  
**Skull, Matteo, 70191 Stuttgart, DE**

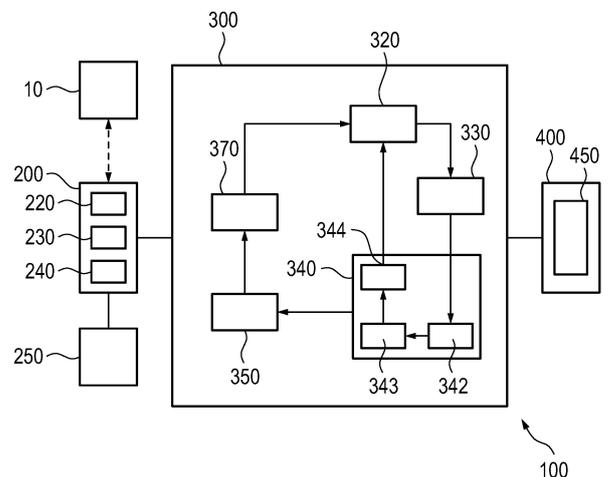
Prüfungsantrag gemäß § 44 PatG ist gestellt.

**Die folgenden Angaben sind den vom Anmelder eingereichten Unterlagen entnommen.**

(54) Bezeichnung: **Verfahren, System und Computerprogrammprodukt zum autonomen Kalibrieren eines elektrischen Antriebsstrangs**

(57) Zusammenfassung: Die Erfindung betrifft Verfahren zum autonomen Kalibrieren eines individuellen elektrischen Antriebsstrangs (10), umfassend:

- Erstellen (S10) eines Trainingsmodells (TM) für einen elektrischen Antriebsstrang (10) von einem Lernverstärkungs-Agenten (320) mittels simulierter Beobachtungen ( $b_1, b_2, \dots, b_n$ ), wobei der Lernverstärkungs-Agent (320) einen Algorithmus des verstärkenden Lernens verwendet;
- Modifizieren (S20) des Trainingsmodells (TM) von dem Lernverstärkungs-Agenten (320) mittels realer Beobachtungen ( $b_{r1}, b_{r2}, \dots, b_{rn}$ ) eines realen idealtypischen Antriebsstrangs (10) zur Erstellung eines simulierten Modells (M), wobei das simulierte Modell (M) Zielzustände ( $sm_1, sm_2, \dots, sm_n$ ) enthält;
- Bestimmen (S30) zumindest eines Zustands ( $s_i$ ) eines individuellen realen elektrischen Antriebsstrangs (10) durch ein Zustandsmodul (350), wobei ein Zustand ( $s_i$ ) durch Parameter ( $p_i$ ) wie Daten und/oder Messwerte von zumindest einer Eigenschaft ( $e_i$ ) des elektrischen Antriebsstrangs (10) definiert wird;
- Übermitteln (S40) des Zustands ( $s_i$ ) an den Lernverstärkungs-Agenten (320);
- Ermitteln (S50) von Kalibrierungsergebnissen (450) für den individuellen realen elektrischen Antriebsstrang (10) von dem Lernverstärkungs-Agenten (320) durch Vergleichen des Zustands ( $s_i$ ) mit zumindest einem Zielzustand ( $sm_{ti}$ ) des simulierten Modells (M).



## Beschreibung

**[0001]** Die Erfindung betrifft ein Verfahren, ein System und ein Computerprogrammprodukt zum autonomen Kalibrieren eines elektrischen Antriebsstrangs.

**[0002]** Die Kalibrierung von Steuervorrichtungen für elektrische Antriebsstränge mit Methoden der künstlichen Intelligenz, insbesondere mit Verfahren des verstärkenden Lernens (engl. reinforcement learning), ist bekannt. Ein elektrischer Antriebsstrang verfügt über einen oder mehrere Elektromotoren, die von einer elektrischen Energieversorgung wie insbesondere einer Batterie oder auch einer Brennstoffzelle mit Energie versorgt werden. Mittels einer Leistungselektronik wie einem Wechselrichter wird die Ausgangsspannung der elektrischen Energieversorgung in Wechselspannung umgewandelt, um den Elektromotor mit der erforderlichen Strom- und Spannungsstärke entsprechend der jeweiligen Antriebssituation zu versorgen. Elektrische Antriebe werden für eine Vielzahl von Funktionen und Aufgaben eingesetzt als Fahrtriebe bei Kraftfahrzeugen, Flurförderfahrzeugen, Bahnen, oder in der Industrie bei Montagemaschinen, oder als Hubantriebe, oder im Bereich der Robotik sowie für eine Vielzahl von Geräten im Haushalt. Ein elektrischer Antrieb hat gegenüber anderen Antriebsarten wie hydraulisch oder pneumatisch den Vorteil einer effizienten Steuerbarkeit des elektrischen Motors als elektromechanischen Energiewandler hinsichtlich des Drehmoments und der Drehzahl durch eine steuerbare Leistungselektronik. Durch Steuerung der Leistungselektronik wird die Leistung des Motors an die jeweilige Aufgabenstellung angepasst, beispielsweise für eine gewünschte Beschleunigung beim Fahren eines Kraftfahrzeugs. Die Steuerung der Leistungselektronik hängt wiederum von der Konfiguration des Motors und anderen Parametern, wie beispielsweise dem Gewicht eines Kraftfahrzeugs ab.

**[0003]** Allerdings wird bei den bekannten Reinforcement-Learning-Verfahren zur Kalibrierung eines elektrischen Antriebsstrangs ein Modell eines real existierenden elektrischen Antriebsstrangs einem Lernverstärkungs-Agenten vorgegeben, das von dem Lernverstärkungs-Agenten nicht verändert wird. Das Modell kann beispielsweise mittels physikalisch-mathematischer Gleichungen beschrieben werden oder auf einer datengesteuerten Grundlage beispielsweise mittels neuronaler Netze erstellt werden. Ein weiterer Ansatz ist die Erstellung eines Modells auf der Basis von Markov-Entscheidungsprozessen. Unabhängig von der Art des zugrundeliegenden Modells für einen elektrischen Antriebsstrang, wird das Modell dem Lernverstärkungs-Agenten jeweils vorgegeben. Dies bedeutet, dass der Lernverstärkungs-Agent das vorgegebene Modell verwendet, um seine Aktionen zu planen.

Der Lernverstärkungs-Agent agiert somit nicht gänzlich unabhängig, da die Auswahl seiner Aktionen von dem verwendeten Modell abhängt.

**[0004]** Das Modell wird üblicherweise von Experten wie Ingenieuren erstellt und bildet eine Umgebung ab, auf die der Lernverstärkungs-Agent zurückgreift. Die Erstellung des Modells, das im Fall eines Antriebsstrangs dessen dynamisches Verhalten beispielsweise hinsichtlich der verwendeten Spannung und der Stromstärke in Abhängigkeit von einer Verkehrssituation widerspiegelt, ist jedoch aufwendig und auch schwierig, so dass das Modell teilweise nicht das tatsächliche Verhalten eines elektrischen Antriebsstrangs abbildet und daher ungenau ist. Zudem ist die Erstellung eines Modells für eine Umgebung innerhalb eines Lernverstärkungsverfahrens mit einem erheblichen Zeit- und damit Kostenaufwand verbunden. Dies führt jedoch dazu, dass auch die Lernergebnisse des Lernverstärkungs-Agenten von dem zugrundeliegenden Modell abhängig sind und daher nur eine eingeschränkte Gültigkeit haben.

**[0005]** Die CN 112632860 A offenbart ein Verfahren zur Identifizierung von Modellparametern eines Energieübertragungssystems, das auf verstärktem Lernen basiert. Das Verstärkungslernverfahren für die Identifikation von Modellparametern des Energieübertragungssystems vermeidet lokale Optimierungen und weist eine hohe Konvergenzrate auf, da es auf einem stufenweisen Identifizierungsprozess beruht.

**[0006]** Die US 2019/0378036 A1 offenbart ein Steuerungsverfahren für Kraftfahrzeuge, welches auf Verfahren des verstärkenden Lernens basiert. Eine Steuereinheit für verstärkendes Lernen wird sowohl auf der Grundlage einer simulierten Bodenfahrzeugumgebung während eines Simulationsmodus und auf der Grundlage einer Kraftfahrzeugumgebung während eines Betriebsmodus eines Kraftfahrzeugs trainiert.

**[0007]** Die DE 10 2019 215 530 A1 offenbart eine Betriebsstrategie für ein Hybridfahrzeug mit einem Elektromotor und einem Verbrennungsmotor, die auf Verfahren des verstärkenden Lernens basiert.

**[0008]** Die DE 10 2019 208 262 A1 offenbart ein Verfahren zum Ermitteln einer Regelungsstrategie für ein technisches System. Die Regelungsstrategie wird basierend auf Modellparameter eines Regelungsmodells erstellt und ausgeführt, wobei zum Auffinden der Regelungsstrategie Reinforcement Learning eingesetzt wird.

**[0009]** Die EP 3 825 263 A1 offenbart ein Verfahren zur computer-implementierten Konfiguration einer regelten Antriebskonfiguration eines Logistiksys-

tems, wobei eine Regelfunktion mittels Reinforcement Learning bestimmt wird.

**[0010]** Die der Erfindung zu Grunde liegende Aufgabe besteht nun darin, ein Verfahren, ein System und ein Computerprogrammprodukt zum autonomen Kalibrieren eines elektrischen Antriebsstrangs zu schaffen, das sich durch eine hohe Zuverlässigkeit, Sicherheit und Genauigkeit auszeichnet und sich einfach implementieren lässt.

**[0011]** Gemäß der vorliegenden Erfindung werden ein Verfahren, ein System und ein Computerprogrammprodukt vorgeschlagen, durch das eine autonome Kalibrierung eines elektrischen Antriebsstrangs ermöglicht wird, um hierdurch die Grundlage für eine zuverlässige und genaue Steuerung des elektrischen Antriebsstrangs zu schaffen.

**[0012]** Diese Aufgabe wird hinsichtlich eines Verfahrens durch die Merkmale des Patentanspruchs 1, hinsichtlich eines Systems durch die Merkmale des Patentanspruchs 10, und hinsichtlich eines Computerprogrammprodukts durch die Merkmale des Patentanspruchs 15 erfindungsgemäß gelöst. Die weiteren Ansprüche betreffen bevorzugte Ausgestaltungen der Erfindung.

**[0013]** Gemäß einem ersten Aspekt stellt die Erfindung ein Verfahren zum autonomen Kalibrieren eines individuellen elektrischen Antriebsstrangs bereit. Das Verfahren umfasst die folgenden Verfahrensschritte:

- Erstellen eines Trainingsmodells für einen elektrischen Antriebsstrang von einem Lernverstärkungs-Agenten mittels simulierter Beobachtungen, wobei der Lernverstärkungs-Agent einen Algorithmus des verstärkenden Lernens verwendet;
- Modifizieren des Trainingsmodells von dem Lernverstärkungs-Agenten mittels realer Beobachtungen eines realen idealtypischen Antriebsstrangs zur Erstellung eines simulierten Modells für den realen idealtypischen elektrischen Antriebsstrang, wobei das simulierte Modell Zielzustände enthält;
- Bestimmen zumindest eines Zustands eines individuellen realen elektrischen Antriebsstrangs durch ein Zustandsmodul, wobei ein Zustand durch Parameter wie Daten und/oder Messwerte von zumindest einer Eigenschaft des elektrischen Antriebsstrangs definiert wird,
- Übermitteln des Zustands an den Lernverstärkungs-Agenten;
- Ermitteln von Kalibrierungsergebnissen für den individuellen realen elektrischen Antriebsstrang von dem Lernverstärkungs-Agenten

durch Vergleichen des Zustands mit zumindest einem Zielzustand des simulierten Modells.

**[0014]** In einer vorteilhaften Ausführungsform ist vorgesehen, dass für das Erstellen eines Trainingsmodells für einen elektrischen Antriebsstrang von einem Lernverstärkungs-Agenten mittels simulierter Beobachtungen ein Umgebungsmodul vorgesehen ist, welches zumindest ein Zustandsuntermodul, ein Belohnungsuntermodul und ein Strategieuntermodul umfasst.

**[0015]** In einer Weiterbildung ist vorgesehen, dass von dem Zustandsuntermodul Zustände generiert werden, die auf den simulierten Beobachtungen beruhen, wobei die simulierten Beobachtungen insbesondere die Stromstärke, die Spannung, das Drehmoment und/oder die Drehzahl eines Elektromotors oder/oder den Ladezustand einer Batterie des elektrischen Antriebsstrangs umfassen.

**[0016]** Insbesondere ist vorgesehen, dass die realen Beobachtungen gemessene Werte von Parametern von einer Eigenschaft des realen idealtypischen Antriebsstrangs darstellen, die von Sensoren ermittelt werden oder in einer Datenbank gespeichert sind.

**[0017]** In einer weiteren Ausführungsform umfasst das Ermitteln von Kalibrierungsergebnissen die folgenden Verfahrensschritte:

- Auswählen einer Berechnungsfunktion und/oder einer Aktion basierend auf einer Richtlinie für einen Zustand für die Modifizierung zumindest eines Parameters von dem Lernverstärkungs-Agenten;
- Berechnen eines modellierten Wertes für die Eigenschaft mittels des modifizierten Parameters;
- Berechnen eines neuen Zustands von einem Umgebungsmodul aufgrund des modellierten Wertes für die Eigenschaft;
- Vergleichen des neuen Zustands mit dem Zielzustand und Zuordnen einer Abweichung für das Vergleichsergebnis in dem Zustandsmodul;
- Ermitteln einer Belohnung von einem Belohnungsmodul für das Vergleichsergebnis;
- Anpassen der Richtlinie des Lernverstärkungs-Agenten basierend auf der Belohnung, wobei bei einer Konvergenz der Richtlinie die optimale Aktion für den berechneten Zustand zurückgegeben wird, und bei einer Nicht-Konvergenz der Richtlinie eine weitere Berechnungsfunktion und/oder eine weitere Aktion für einen Zustand mit einer Modifizierung zumindest eines Parameters von dem Lernverstärkungs-Agenten

ausgewählt wird, solange bis der Ziel-Zustand erreicht ist.

**[0018]** Insbesondere sind Sensoren und/oder Messvorrichtungen zur Bestimmung der Parameter von den Eigenschaften des individuellen elektrischen Antriebsstrangs vorgesehen.

**[0019]** Vorteilhafterweise sind eine positive Aktion A+, die den Wert für einen Parameter erhöht, eine neutrale Aktion A0, bei der der Wert des Parameters gleichbleibt, und eine negative Aktion A-, bei der sich der Wert des Parameters verringert, vorgesehen sind.

**[0020]** In einer Ausführungsform ist vorgesehen, dass das Belohnungsmodul eine Datenbank oder Matrix für die Bewertung der Aktionen umfasst.

**[0021]** Insbesondere ist der zumindest eine Algorithmus des Lernverstärkungs-Agenten als Markow-Entscheidungsprozess, Temporal Difference Learning (TD-Learning), Q-Learning, SARSA, Monte-Carlo-Simulation oder Actor-Critic ausgebildet.

**[0022]** Gemäß einem zweiten Aspekt stellt die Erfindung ein System zum autonomen Kalibrieren eines individuellen elektrischen Antriebsstrangs bereit. Das System umfasst ein Eingabemodul, ein Lernverstärkungsmodul und ein Ausgabemodul. Das Lernverstärkungsmodul umfasst einen Lernverstärkungs-Agenten, der einen Algorithmus des verstärkenden Lernens verwendet, ein Aktionsmodul, ein Umgebungsmodul, ein Zustandsmodul und ein Belohnungsmodul. Der Lernverstärkungs-Agent ist ausgebildet, ein Trainingsmodell für einen elektrischen Antriebsstrang mittels simulierter Beobachtungen zu erstellen, und das Trainingsmodell mittels realer Beobachtungen eines realen idealtypischen Antriebsstrangs zu modifizieren zur Erstellung eines simulierten Modells für den realen idealtypischen elektrischen Antriebsstrang, wobei das simulierte Modell Zielzustände enthält. Das Zustandsmodul ist ausgebildet, zumindest einen Zustand eines individuellen realen elektrischen Antriebsstrangs zu bestimmen, wobei ein Zustand durch Parameter wie Daten und/oder Messwerte von zumindest einer Eigenschaft des elektrischen Antriebsstrangs definiert wird, und den Zustand an den Lernverstärkungs-Agenten zu übermitteln. Der Lernverstärkungs-Agent ist ausgebildet, Kalibrierungsergebnisse für den individuellen realen elektrischen Antriebsstrang durch Vergleichen des Zustands mit zumindest einem Zielzustand des simulierten Modells zu ermitteln.

**[0023]** In einer Weiterentwicklung ist vorgesehen, dass das Umgebungsmodul zumindest ein Zustandsuntermodul, ein Belohnungsuntermodul und ein Strategieuntermodul umfasst.

**[0024]** In einer weiteren Ausführungsform ist vorgesehen, dass das Zustandsuntermodul ausgebildet ist, Zustände zu generieren, die auf den simulierten Beobachtungen beruhen, wobei die simulierten Beobachtungen insbesondere die Stromstärke, die Spannung, das Drehmoment und/oder die Drehzahl eines Elektromotors und/oder den Ladezustand einer Batterie des elektrischen Antriebsstrangs umfassen.

**[0025]** Insbesondere ist vorgesehen, dass die realen Beobachtungen gemessene Werte von Parametern von einer Eigenschaft darstellen, die von Sensoren ermittelt werden oder die in einer Datenbank gespeichert sind.

**[0026]** Vorteilhafterweise sind Sensoren und/oder Messvorrichtungen zur Bestimmung der Parameter von den Eigenschaften des individuellen elektrischen Antriebsstrangs vorgesehen sind.

**[0027]** Gemäß einem dritten Aspekt stellt die Erfindung ein Computerprogrammprodukt bereit, das einen ausführbaren Programmcode umfasst, der so konfiguriert ist, dass er bei seiner Ausführung das Verfahren gemäß dem ersten Aspekt ausführt.

**[0028]** Nachfolgend wird die Erfindung anhand von in der Zeichnung dargestellten Ausführungsbeispielen näher erläutert.

**[0029]** Dabei zeigt:

**Fig. 1** ein Blockdiagramm zur Erläuterung eines Ausführungsbeispiels eines erfindungsgemäßen Systems;

**Fig. 2** ein Flussdiagramm zur Erläuterung der einzelnen Verfahrensschritte eines erfindungsgemäßen Verfahrens;

**Fig. 3** ein Blockdiagramm eines Computerprogrammprodukt gemäß einer Ausführungsform des dritten Aspekts der Erfindung.

**[0030]** Zusätzliche Kennzeichen, Aspekte und Vorteile der Erfindung oder ihrer Ausführungsbeispiele werden durch die ausführliche Beschreibung in Verbindung mit den Ansprüchen ersichtlich.

**[0031]** Fig. 1 zeigt ein erfindungsgemäßes System 100 zum autonomen Kalibrieren eines elektrischen Antriebsstrangs 10. Ein elektrischer Antriebsstrang 10 verfügt über einen oder mehrere Elektromotoren, die von einer elektrischen Energieversorgung wie insbesondere einer Batterie oder auch einer Brennstoffzelle mit Energie versorgt werden. Mittels einer Leistungselektronik wie einem Wechselrichter wird die Ausgangsspannung der elektrischen Energieversorgung in Wechselspannung umgewandelt, um den Elektromotor mit der erforderlichen Strom- und Spannungsstärke entsprechend der jeweiligen Antriebs-

tuation zu versorgen. Elektrische Antriebe werden für eine Vielzahl von Funktionen und Aufgaben eingesetzt als Fahrtriebe bei Kraftfahrzeugen, Flurförderfahrzeugen, Bahnen, oder in der Industrie bei Montagemaschinen, oder als Hubantriebe, oder im Bereich der Robotik sowie für eine Vielzahl von Geräten im Haushalt. Ein elektrischer Antrieb hat gegenüber anderen Antriebsarten wie hydraulisch oder pneumatisch den Vorteil einer effizienten Steuerbarkeit des elektrischen Motors als elektromechanischen Energiewandler hinsichtlich des Drehmoments und der Drehzahl durch eine steuerbare Leistungselektronik. Durch Steuerung der Leistungselektronik an die jeweilige Aufgabenstellung wird die Leistung des Motors angepasst, beispielsweise für eine gewünschte Beschleunigung beim Fahren eines Kraftfahrzeugs. Die Steuerung der Leistungselektronik hängt wiederum von der Konfiguration des Motors und anderen Parametern, wie beispielsweise dem Gewicht eines Kraftfahrzeugs ab.

**[0032]** Das erfindungsgemäße System 100 beruht auf Methoden des verstärkenden Lernens (engl. reinforcement learning) und umfasst ein Eingabemodul 200, ein Lernverstärkungsmodul 300 und ein Ausgabemodul 400. Das Lernverstärkungsmodul 300 umfasst einen Lernverstärkungs-Agenten (LV-Agenten) 320, ein Aktionsmodul 330, ein Umgebungsmodul 340, ein Zustandsmodul 350 und ein Belohnungsmodul 370.

**[0033]** Das Eingabemodul 200, das Lernverstärkungsmodul 300 und das Ausgabemodul 400 können jeweils mit einem Prozessor und/oder einer Speichereinheit versehen sein.

**[0034]** Unter einem „Prozessor“ kann im Zusammenhang mit der Erfindung beispielsweise eine Maschine oder eine elektronische Schaltung verstanden werden. Bei einem Prozessor kann es sich insbesondere um einen Hauptprozessor (engl. Central Processing Unit, CPU), einen Mikroprozessor oder einen Mikrocontroller, beispielsweise eine anwendungsspezifische integrierte Schaltung oder einen digitalen Signalprozessor, möglicherweise in Kombination mit einer Speichereinheit zum Speichern von Programmbefehlen, etc. handeln. Auch kann unter einem Prozessor ein virtualisierter Prozessor, eine virtuelle Maschine oder eine Soft-CPU verstanden werden. Es kann sich beispielsweise auch um einen programmierbaren Prozessor handeln, der mit Konfigurationsschritten zur Ausführung des genannten erfindungsgemäßen Verfahrens ausgerüstet wird oder mit Konfigurationsschritten derart konfiguriert ist, dass der programmierbare Prozessor die erfindungsgemäßen Merkmale des Verfahrens, der Komponente, der Module, oder anderer Aspekte und/oder Teilaspekte der Erfindung realisiert.

**[0035]** Unter einer „Speichereinheit“ oder einem „Speichermodul“ und dergleichen kann im Zusammenhang mit der Erfindung beispielsweise ein flüchtiger Speicher in Form eines Arbeitsspeichers (engl. Random-Access Memory, RAM) oder ein dauerhafter Speicher wie eine Festplatte oder ein Datenträger oder beispielsweise ein wechselbares Speichermodul verstanden werden. Es kann sich bei dem Speichermodul aber auch um eine cloudbasierte Speicherlösung handeln.

**[0036]** Unter einem „Modul“ kann im Zusammenhang mit der Erfindung beispielsweise ein Prozessor und/oder eine Speichereinheit zum Speichern von Programmbefehlen verstanden werden. Beispielsweise ist der Prozessor speziell dazu eingerichtet, die Programmbefehle derart auszuführen, damit der Prozessor und/oder die Steuereinheit Funktionen ausführt, um das erfindungsgemäße Verfahren oder einen Schritt des erfindungsgemäßen Verfahrens zu implementieren oder realisieren.

**[0037]** Unter „Daten“ sind im Zusammenhang mit der Erfindung sowohl Rohdaten als auch bereits aufbereitete Daten beispielsweise aus Messergebnissen von Sensoren oder aus Simulationsergebnissen zu verstehen.

**[0038]** Das verstärkende Lernen beruht darauf, dass der LV-Agent 320 für einen bestimmten Zustand  $s_i \in S$  aus einer Menge verfügbarer Zustände zumindest eine Aktion  $a_i \in A$  aus einer Menge verfügbarer Aktionen auswählt. Die Wahl der ausgewählten Aktion  $a_i$  basiert auf einer Strategie bzw. Richtlinie (engl.: policy). Für die ausgewählte Aktion  $a_i$  erhält der LV-Agent 320 eine Belohnung (engl.: reward)  $r_i \in R$  von dem Belohnungsmodul 370. Die Zustände  $s_i \in S$  erhält der Agent 320 von dem Zustandsmodul 350, auf das der LV-Agent 320 zugreifen kann. Die Strategie wird auf der Grundlage der erhaltenen Belohnungen  $r_i$  von dem LV-Agenten 320 angepasst. In der Strategie ist festgelegt, welche Aktion  $a_i \in A$  aus der Menge verfügbarer Aktionen für einen bestimmten Zustand  $s_i \in S$  aus der Menge verfügbarer Zustände ausgewählt werden soll. Hierdurch wird ein neuer Zustand  $s_{i+1}$  erzeugt, für den der LV-Agent 320 eine Belohnung  $r_{i+1}$  erhält. Eine Strategie legt somit die Zuordnung zwischen einem Zustand  $s_i$  und einer Aktion  $a_i$  fest, so dass die Strategie die Wahl der auszuführenden Aktion  $a_i$  für einen Zustand  $s_i$  angibt. Das Ziel des LV-Agenten 320 besteht darin, die erzielten Belohnungen  $r_i, r_{i+1}, \dots, r_{i+n}$  zu maximieren.

**[0039]** Im Aktionsmodul 330 werden die vom LV-Agenten 320 ausgewählten Aktionen  $a_i$  durchgeführt. Durch eine Aktion  $a_i$  wird beispielsweise eine Anpassung eines Wertes eines Parameters  $p_i \in P$  aus der Menge von Parametern für zumindest eine Eigenschaft  $e_i$  einer technischen Komponente des elektri-

schen Antriebsstrangs vorgenommen. Vorzugsweise handelt es sich bei der Aktion  $a_i$  um eine der Aktionen  $A(+)$ ,  $A(0)$  und  $A(-)$ . Bei einer positiven Aktion  $A(+)$  handelt es sich um eine Aktion, die den Wert für einen Parameter  $p_i$  erhöht, bei einer neutralen Aktion  $A(0)$  handelt es sich um ein Aktion, bei der der Wert des Parameters  $p_i$  gleichbleibt, während sich bei einer negativen Aktion  $A(-)$  der Wert des Parameters  $p_i$  verringert.

**[0040]** Das Umgebungsmodul 340 berechnet aufgrund der gewählten Aktion  $a_i$  und unter Berücksichtigung von zuvor festgelegten Randbedingungen (engl.: constraints) die Zustände  $s_i \in S$ . Bei den Randbedingungen kann es sich auch um wirtschaftliche Aspekte wie die Kostenstruktur, die Energiekosten, die Umweltbilanz, die Verfügbarkeit oder die Liefersituation handeln.

**[0041]** Ein Zustand  $s_i \in S$  ist somit durch die Auswahl von bestimmten Werten für Parameter  $p_i$  von Eigenschaften  $e_i$  des elektrischen Antriebsstrangs 10 definiert. Bei den Eigenschaften  $e_i$  kann es sich beispielsweise um ein Spannungsverhalten, einen elektrischen Widerstand oder eine Kennlinie zum Drehmoment/Drehzahl-Verhalten eines Elektromotors des elektrischen Antriebsstrangs handeln. Ein Parameterwert  $p_i$  gibt die konkrete Spannung oder das konkrete Drehmoment für diese Eigenschaft  $e_i$  wieder.

**[0042]** In dem Zustandsmodul 350 wird eine Abweichung  $\Delta$  zwischen einem Zielzustand  $s_t$  und dem berechneten Zustand  $s_i$  berechnet. Der Endzustand ist erreicht, wenn die berechneten Zustände  $s_i$  gleich oder größer als die Ziel-Zustände  $s_t$  sind.

**[0043]** In dem Belohnungsmodul 370 wird dem Grad der Abweichung  $\Delta$  zwischen dem berechneten Wert für den Zustand  $s_i$  und dem Zielwert des Zustands  $s_t$  eine Belohnung  $r_i$  zugeordnet. Da der Grad der Abweichung  $\Delta$  von der Auswahl der jeweiligen Aktion  $A(+)$ ,  $A(0)$ ,  $A(-)$  abhängig ist, wird vorzugsweise in einer Matrix oder einer Datenbank der jeweiligen ausgewählten Aktion  $A(+)$ ,  $A(0)$ ,  $A(-)$  die Belohnung  $r_i$  zugeordnet. Eine Belohnung  $r_i$  weist vorzugsweise die Werte +1 und -1 auf, wobei eine geringe oder positive Abweichung  $\Delta$  zwischen dem berechneten Zustand  $s_i$  und dem Zielzustand  $s_t$  mit +1 belohnt wird und somit verstärkt wird, während eine erhebliche negative Abweichung  $\Delta$  mit -1 belohnt wird und somit negativ bewertet wird. Es ist aber auch denkbar, dass Werte  $> 1$  und Werte  $< 1$  Verwendung finden.

**[0044]** Vorzugsweise wird als Algorithmus für den LV-Agenten 320 ein Markov-Entscheidungsprozess verwendet. Es kann aber auch vorgesehen sind, einen Temporal Difference Learning (TD-Learning) Algorithmus zu verwenden. Ein LV-Agent 320 mit

einem TD-Learning-Algorithmus nimmt die Anpassung der Aktionen  $A(+)$ ,  $A(0)$ ,  $A(-)$  nicht erst dann vor, wenn er die Belohnung erhält, sondern nach jeder Aktion  $a_i$  auf Basis einer geschätzten erwarteten Belohnung. Des Weiteren sind auch Algorithmen wie Q-Learning und SARSA denkbar oder Actor-Critic oder auch Monte-Carlo-Simulationen. Mit diesen Algorithmen ist eine dynamische Programmierung und eine Anpassung der Strategie durch Iterationsverfahren möglich.

**[0045]** Darüber hinaus enthält/enthalten der LV-Agent 320 und/oder das Aktionsmodul 330 und/oder das Umgebungsmodul 340 und/oder das Zustandsmodul 350 und/oder das Belohnungsmodul 370 Berechnungsverfahren und Algorithmen  $f_i$  für mathematische Regressionsverfahren oder physikalische Modellberechnungen, die eine Korrelation zwischen ausgewählten Parametern  $p_i \in P$  aus einer Menge von Parametern und den Ziel-Zuständen  $s_t$  beschreiben. Bei den mathematischen Funktionen  $f_i$  kann es sich um statistische Verfahren wie Mittelwerte, Minimal- und Maximalwerte, Lookup Tabellen, Modelle zu Erwartungswerten, lineare Regressionsverfahren oder Gauß-Prozesse handeln, um Fast Fourier Transformationen, Integral- und Differentialrechnungen, Markov-Verfahren, Wahrscheinlichkeitsverfahren wie Monte Carlo-Verfahren, Temporal Difference Learning, aber auch um erweiterte Kalman-Filter, radiale Basisfunktionen, Datenfelder, oder auch um konvergente neuronale Netze, tiefe neuronale Netze, rückgekoppelte/rekurrente neuronale Netze oder gefaltete Netze (engl. convolutional neural network) handeln. Basierend auf den Aktionen  $a_i$  und den Belohnungen  $r_i$  wählt/wählen der LV-Agent 320 und/oder das Aktions-Modul 330 und/oder das Umgebungsmodul 340 und/oder das Zustandsmodul 350 und/oder das Belohnungsmodul 370 für einen Zustand  $s_i$  eine oder mehrere dieser Berechnungsfunktionen  $f_i$  aus.

**[0046]** Ein neuronales Netzwerk besteht aus Neuronen, die in mehreren Schichten angeordnet und unterschiedlich miteinander verbunden sind. Ein Neuron ist in der Lage, an seinem Eingang Informationen von außerhalb oder von einem anderen Neuron entgegenzunehmen, die Information in einer bestimmten Art zu bewerten und sie in veränderter Form am Neuronen-Ausgang an ein weiteres Neuron weiterzuleiten oder als Endergebnis auszugeben. Hidden-Neuronen sind zwischen den Input-Neuronen und Output-Neuronen angeordnet. Je nach Netzwerktyp können mehrere Schichten von Hidden-Neuronen vorhanden sein. Sie sorgen für die Weiterleitung und Verarbeitung der Informationen. Output-Neuronen liefern schließlich ein Ergebnis und geben dieses an die Außenwelt aus. Durch die Anordnung und die Verknüpfung der Neuronen entstehen verschiedene Typen von neuronalen Netzwerken wie Feedforward-Netzwerke, Rekurrente

Netzwerke oder gefaltete neuronale Netze (engl. Convolutional Neural Networks). Ein gefaltetes neuronales Netzwerk (engl. Convolutional Neural Network) besitzt mehrere Faltungsschichten und ist für maschinelles Lernen und Anwendungen mit Künstlicher Intelligenz (KI) im Bereich der Mustererkennung sehr gut geeignet. Die Netzwerke lassen sich durch unbeaufsichtigtes oder überwachtes Lernen trainieren.

**[0047]** Während in einem klassischen Umgebungsmodul 340 ein Modell eines elektrischen Antriebsstrangs 10 vorgegeben ist, welches die Zielzustände  $s_{t1}, s_{t2}, \dots, s_{tn}$  enthält, entwickelt der Lernverstärkungs-Agent 320 gemäß der vorliegenden Erfindung das Modell des elektrischen Antriebsstrangs 10 selbstständig und autonom. Dabei wird das Modell des elektrischen Antriebsstrangs 10 durch eine Vielzahl von Aktionen  $a_i \in A$  von dem Lernverstärkungs-Agenten 320 gelernt und bildet dann die Basis für die Kalibrierung eines realen elektrischen Antriebsstrangs 10 durch das Lernverstärkungsmodul 300.

**[0048]** Der erfindungsgemäße Gedanke besteht somit darin, einen realen elektrischen Antriebsstrang 10 mit Hilfe eines modellbasierten verstärkenden Lernens zu kalibrieren, bei dem das Modell des elektrischen Antriebsstrangs 10 nicht vorliegen muss, sondern von dem LV-Agenten 320 selbst modelliert wird. Dabei simuliert das vom LV-Agenten 320 erstellte Modell des elektrischen Antriebsstrangs 10 nicht im Detail die Physik oder Dynamik des elektrischen Antriebsstrangs 10, sondern das Modell wird mittels einer Vielzahl von Interaktionen zwischen Aktionen, Zuständen und Belohnungen, die vom LV-Agenten 320 ausgeführt werden, entwickelt. Die Fragestellung des LV-Agenten 320 lautet daher immer, welche Zustände es gibt und was passiert, wenn er für einen bestimmten Zustand eine Aktion ausführt, und wie die Belohnung aussieht, wenn er für diesen bestimmten Zustand eine Aktion ausführt.

**[0049]** Um ein Modell eines elektrischen Antriebsstrangs 10 zu erstellen, ist erfindungsgemäß vorgesehen, dass das Umgebungsmodul 340 zumindest drei Untermodule aufweist. Das erste Untermodul ist als Zustandsuntermodul 342, das zweite Untermodul als Belohnungsuntermodul 343 und das dritte Untermodul ist als Strategieuntermodul 344 ausgebildet.

**[0050]** Das Zustandsuntermodul 342 repräsentiert Zustände  $su_1, su_2, \dots, su_n$ , die der LV-Agent 320 auswählen kann, wobei der ausgewählte Zustand  $su_j$  dann der Zustand ist, in dem sich der LV-Agent 320 aktuell befindet. Ein Zustand  $su_j$  wird simuliert und beruht auf simulierten Beobachtungen  $b_1, b_2, \dots, b_n$ , die dem Zustandsuntermodul 342 in Form von Eingangsdaten 220 von dem Eingabemodul 200 zugeführt werden. Der LV-Agent 320 lernt die Zustände  $su_1, su_2, \dots, su_n$  des Zustandsuntermoduls 342

durch das Sammeln der Beobachtungen  $b_1, b_2, \dots, b_n$ . Für die gesammelten Beobachtungen  $b_1, b_2, \dots, b_n$  entwirft er ein Modell, das die Zustände  $su_1, su_2, \dots, su_n$ , in denen er sich befinden kann, definiert und das eine Funktion der gesammelten Beobachtungen  $b_1, b_2, \dots, b_n$  darstellt. Für die Entwicklung des Modells verwendet er insbesondere neuronale Netze. Bei den Beobachtungen  $b_1, b_2, \dots, b_n$  kann es sich beispielsweise um die Stromstärke, die Spannung, das Drehmoment und die Drehzahl eines Elektromotors oder den Ladezustand einer Batterie des elektrischen Antriebsstrangs 10 handeln. Mögliche Zustände  $su_1, su_2, \dots, su_n$  des Zustandsuntermoduls 342 leiten sich somit aus diesen simulierten Beobachtungen  $b_1, b_2, \dots, b_n$ , wie einem Drehmoment oder einer Drehzahl eines Elektromotors, ab.

**[0051]** Das Belohnungsuntermodul 343 ordnet den ermittelten Zuständen  $su_1, su_2, \dots, su_n$  Belohnungen  $ru_1, ru_2, \dots, ru_n$  zu.

**[0052]** Das Strategieuntermodul 344 entwickelt eine Strategie zur Ermittlung von neuen Zuständen  $su_{1+1}, su_{2+1}, \dots, su_{n+1}$ , indem es vorschlägt, welche Aktionen  $a_j$  von den  $a_1, a_2, \dots, a_n$  Aktionen aus dem Aktionsuntermodul 330 auf die alten Zustände  $su_1, su_2, \dots, su_n$  angewendet werden sollen. Durch die Anwendung der mittels der Strategie ausgewählten Aktionen  $a_1, a_2, \dots, a_n$  werden neue Zustände  $su_{1+1}, su_{2+1}, \dots, su_{n+1}$  generiert, die dann wieder dem Zustandsuntermodul 342 zugeführt werden. In dem Belohnungsuntermodul 343 werden den neu ermittelten Zuständen  $su_{1+1}, su_{2+1}, \dots, su_{n+1}$  wiederum Belohnungen  $ru_{1+1}, ru_{2+1}, \dots, ru_{n+1}$  zugeordnet.

**[0053]** Das Umgebungsmodul 340 führt die Berechnungen so lange durch, bis sich ein stabiles Zustandsniveau eingependelt hat. Dieses Zustandsniveau kann einen Zielzustand  $su_{tj}$  oder eine Vielzahl von Zielzuständen  $su_{t1}, su_{t2}, \dots, su_{tn}$  für den LV-Agenten 320 umfassen. Das Ergebnis des Umgebungsmoduls 340 besteht somit in den berechneten Zielzuständen  $su_{t1}, su_{t2}, \dots, su_{tn}$ , die ein Trainingsmodell TM des elektrischen Antriebsstrangs 10 darstellen.

**[0054]** Für die Trainingsphase werden an das Zustandsuntermodul 342 von dem Eingabemodul 200 somit beliebige oder ausgewählte simulierte Beobachtungen  $b_1, b_2, \dots, b_n$  als Eingangsdaten 220 übermittelt. Aus diesen Eingangsdaten 220 entwickelt der LV-Agent 320 autonom ein erstes Trainingsmodell TM des elektrischen Antriebsstrangs 10. Dieses Modell wird durch die Zielzustände  $su_{t1}, su_{t2}, \dots, su_{tn}$  und die angewandte Strategie beschrieben.

**[0055]** An die Trainingsphase schließt sich die Modellierungsphase an, in der das Trainingsmodell TM in ein Modell M eines realen elektrischen Antriebsstrangs 10 transformiert wird. Bei dem rea-

len elektrischen Antriebsstrang 10 handelt es sich um eine idealtypische Ausprägung, bei der eine gewünschte Dynamik beispielsweise hinsichtlich des Verhältnisses von Drehmoment und Drehzahl gegeben ist. In der Modellierungsphase werden dem Zustandsuntermodul 342 von dem Eingabemodul 200 reale Beobachtungen  $b_{r1}, b_{r2}, \dots, b_{rn}$  als Daten 230 zugeführt, aus denen die realen Zustände  $su_{r1}, su_{r2}, \dots, su_{rn}$  generiert werden. Die realen Beobachtungen  $b_{r1}, b_{r2}, \dots, b_{rn}$  können gemessene Parameterwerte  $p_i$  von einer Eigenschaft  $e_i$  darstellen, die von hier nicht näher beschriebenen Sensoren ermittelt worden sind. Vorzugsweise sind die Parameterwerte in einer Datenbank 250 gespeichert, die mit dem Eingabemodul 200 in Verbindung steht.

**[0056]** In dem Belohnungsmodul 343 wird nun eine Abweichung  $\Delta$  zwischen den realen Zuständen  $su_{r1}, su_{r2}, \dots, su_{rn}$  und den während der Trainingsphase generierten Zielzuständen  $su_{t1}, su_{t2}, \dots, su_{tn}$  berechnet. Zudem wird in dem Belohnungsmodul 343 der Grad der Abweichung  $\Delta$  zwischen dem realen Zustand  $su_{ri}$  und dem Zielwert des Zielzustands  $su_{ti}$  eine Belohnung  $r_{i+1}$  zugeordnet.

**[0057]** Das Strategieuntermodul 344 entwickelt aufgrund der neuen Belohnungen  $r_{1+1}, r_{2+1}, \dots, r_{n+1}$  eine geänderte Strategie zur Ermittlung von neuen Zuständen  $su_{1+1}, su_{2+1}, \dots, su_{n+1}$ , indem es vorschlägt, welche Aktionen  $a_j$  von den  $a_1, a_2, \dots, a_n$  aus dem Aktionsuntermodul 330 auf die alten Zielzustände  $su_{t1}, su_{t2}, \dots, su_{tn}$  angewendet werden sollen. Der Endzustand ist erreicht, wenn die generierten Zustände  $su_{t1+1}, su_{t2+1}, \dots, su_{tn+1}$  gleich oder größer als die realen Zustände  $su_{r1}, su_{r2}, \dots, su_{rn}$  sind, da dann das Trainingsmodell TM in ein Modell M transformiert wurde, das einen realen idealtypischen elektrischen Antriebsstrang wiedergibt.

**[0058]** Dieses Modell M eines realen elektrischen Antriebsstrang 10 stellt nun die Zielzustände  $sm_{t1}, sm_{t2}, \dots, sm_{tn}$  zur Verfügung, mit denen eine Kalibrierung eines individuellen realen elektrischen Antriebsstrangs 10 durch den LV-Agenten 320 durchgeführt werden kann.

**[0059]** Hierzu werden dem Zustandsmodul 350 Werte von Parametern  $p_i$  von Eigenschaften  $e_i$  eines individuellen elektrischen Antriebsstrangs 10 von dem Eingabemodul 200 in Form von realen Daten 240 zugeführt. Die Parameterwerte  $p_i$  können durch hier nicht näher beschriebene Sensoren gemessen werden. Bei den Sensoren handelt es sich insbesondere um Drucksensoren, Drehmomentensensoren, Drehzahlsensoren, Beschleunigungssensoren, Geschwindigkeitssensoren, kapazitive Sensoren, induktive Sensoren und Temperatursensoren.

**[0060]** Ein Zustand  $s_i \in S$  eines individuellen elektrischen Antriebsstrangs 10 ist somit durch die Aus-

wahl von Werten von Parametern  $p_i$  von Eigenschaften  $e_i$  definiert. Bei den Eigenschaften  $e_i$  kann es sich beispielsweise um ein Spannungsverhalten, einen elektrischen Widerstand oder eine Kennlinie zum Drehmoment/Drehzahl-Verhalten eines Elektromotors des elektrischen Antriebsstrangs handeln. Ein Parameterwert  $p_i$  gibt die konkrete Spannung oder das konkrete Drehmoment für diese Eigenschaft  $e_i$  wieder.

**[0061]** Der LV-Agent wählt für diese Zustände  $s_1, s_2, \dots, s_n$  wie im Obigen beschriebenen Aktionen (A+), (A0) und (A-) aus, um eine Anpassung an die Zielzustände  $sm_{t1}, sm_{t2}, \dots, sm_{tn}$  des generierten Modells M vorzunehmen. Das Umgebungsmodul 340 berechnet aufgrund der gewählten Aktion  $a_i$  und unter Berücksichtigung von zuvor festgelegten Randbedingungen (engl.: constraints) neue Zustände  $s_{i+1} \in S$ . Bei den Randbedingungen kann es sich auch um wirtschaftliche Aspekte wie die Kostenstruktur, die Energiekosten, die Umweltbilanz, die Verfügbarkeit oder die Liefersituation handeln.

**[0062]** In dem Zustandsmodul 350 wird eine Abweichung  $\Delta$  zwischen einem Zielzustand  $s_t$  und dem berechneten Zustand  $s_{i+1}$  berechnet. In dem Belohnungsmodul 370 wird dem Grad der Abweichung  $\Delta$  zwischen dem berechneten Wert für den Zustand  $s_{i+1}$  und dem Zielwert des Zustands  $sm_t$  eine Belohnung  $r_i$  zugeordnet.

**[0063]** Dann beginnt ein zweiter Zyklus, bei dem der LV-Agent 320 eine andere Aktion  $a_{i+1}$  und/oder eine andere Berechnungsfunktion  $f_{i+1}$  und/oder einen anderen Parameter  $p_{i+1}$  ausgewählt entsprechend der definierten Strategie bzw. Richtlinie. Das Ergebnis wird wiederum dem Zustandsmodul 350 zugeführt und das Ergebnis des Vergleichs im Belohnungsmodul 370 bewertet. Der LV-Agent 320 wiederholt den Kalibrierungsvorgang für alle vorgesehenen Aktionen  $a_i, a_{i+1}, \dots, a_{i+n}$ , Berechnungsfunktionen  $f_i, f_{i+1}, \dots, f_{i+n}$  und Parameter  $p_i, p_{i+1}, \dots, p_{i+n}$  solange, bis eine größtmögliche Übereinstimmung zwischen einem berechneten Zustand  $s_{i+1}, s_{i+2}, \dots, s_{i+n}$  und einem Ziel-Zustand  $sm_{ti}$  erreicht ist. Vorzugsweise ist der Endzustand der Kalibrierung erreicht, wenn die Abweichung  $\Delta$  im Bereich von +/-5% liegt. Der LV-Agent 320 optimiert somit sein Verhalten und damit die Strategie, nach der eine Aktion  $a_i$  ausgewählt wird, solange, bis die berechneten Zustände  $s_{i+1}, s_{i+2}, \dots, s_{i+n}$  konvergieren. Der Endzustand ist erreicht, wenn die berechneten Zustände  $s_{i+1}, s_{i+2}, \dots, s_{i+n}$  gleich oder größer als die Ziel-Zustände  $sm_1, sm_2, \dots, sm_n$  sind. Das Kalibrierungsergebnis kann in Form von Ausgabedaten 450 auf dem Ausgabemodul 400 ausgegeben werden. Das Eingabemodul 200 und das Ausgabemodul 400 können in einem Hardwaregerät wie einem Computer, einem Tablet, einem Smartphone, etc. integriert sein.

**[0064]** Insbesondere kann vorgesehen sein, dass die Berechnungsergebnisse in Form von Zuständen, Aktionen, Belohnungen und Strategien in einer Cloud-Computing-Infrastruktur gespeichert werden und jeweils über das Internet verfügbar sind. Der LV-Agent 320, das Aktionsmodul 330, das Umgebungsmodul 340, das Zustandsmodul 350 und das Belohnungsmodul 370 verfügen hierzu über die erforderlichen technischen Schnittstellen und Protokolle für den Zugang zu der Cloud-Computing-Infrastruktur. Hierdurch kann die Recheneffizienz gesteigert werden, da die Zugriffsmöglichkeiten und die Zugriffsgeschwindigkeiten zu bereits berechneten Zuständen, Aktionen, Belohnungen und Strategien vereinfacht ist.

**[0065]** In Fig. 2 sind die Verfahrensschritte zum autonomen Kalibrieren eines individuellen elektrischen Antriebsstrangs 10 dargestellt.

**[0066]** In einem Schritt S10 wird ein Trainingsmodells TM für einen elektrischen Antriebsstrangs 10 von einem Lernverstärkungs-Agenten 320 mittels simulierter Beobachtungen  $b_1, b_2, \dots, b_n$  erstellt, wobei der Lernverstärkungs-Agent 320 einen Algorithmus des verstärkenden Lernens verwendet.

**[0067]** In einem Schritt S20 wird das Trainingsmodell TM von dem Lernverstärkungs-Agenten 320 modifiziert mittels realer Beobachtungen  $b_{r1}, b_{r2}, \dots, b_{rn}$  eines realen idealtypischen Antriebsstrangs 10 zur Erstellung eines simulierten Modells M für den realen idealtypischen elektrischen Antriebsstrang 10, wobei das simulierte Modell M Zielzustände  $sm_1, sm_2, \dots, sm_n$  enthält.

**[0068]** In einem Schritt S30 wird zumindest ein Zustand  $s_i$  eines individuellen realen elektrischen Antriebsstrangs 10 durch ein Zustandsmodul 350 bestimmt, wobei ein Zustand  $s_i$  durch Parameter  $p_i$  wie Daten und/oder Messwerte von zumindest einer Eigenschaft  $e_i$  des elektrischen Antriebsstrangs 10 definiert wird.

**[0069]** In einem Schritt S40 wird der Zustand  $s_i$  an den Lernverstärkungs-Agenten 320 übermittelt.

**[0070]** In einem Schritt S50 werden Kalibrierungsergebnisse 450 für den individuellen realen elektrischen Antriebsstrang 10 von dem Lernverstärkungs-Agenten 320 durch Vergleichen des Zustands  $s_i$  mit zumindest einem Zielzustand  $sm_{ij}$  des simulierten Modells M ermittelt.

**[0071]** Fig. 3 stellt schematisch ein Computerprogrammprodukt 900 dar, das einen ausführbaren Programmcode 950 umfasst, der konfiguriert ist, um das Verfahren gemäß dem ersten Aspekt der vorliegenden Erfindung auszuführen, wenn es ausgeführt wird.

**[0072]** Mit dem Verfahren und dem System 100 gemäß der vorliegenden Erfindung kann somit zuverlässig ein elektrischer Antriebsstrang 10 mittels Methoden des verstärkenden Lernens kalibriert werden, ohne dass ein detailliertes Modell eines realen elektrischen Antriebsstrangs 10 in dem Umgebungsmodul 340 des Lernverstärkungsmoduls 300 vorliegen muss. Die Modellierung eines realen elektrischen Antriebsstrangs wird vielmehr von dem LV-Agenten 320 selbstständig und autonom durchgeführt. Dadurch werden die Zielzustände, die bei einer Kalibrierung erreicht werden sollen, durch das von dem LV-Agenten erstellte Modell vorgegeben. Die Zielzustände sind genauer und ermöglichen daher eine verbesserte Kalibrierung. Mit der vorliegenden Erfindung kann somit in kurzer Zeit und mit verringerten Kosten eine zuverlässige Kalibrierung von elektrischen Antriebsträngen durchgeführt werden.

#### Bezugszeichenliste

10	elektrischer Antriebsstrang
100	System
200	Eingabemodul
220	simulierte Daten
230	reale Daten
240	reale Daten
250	Datenbank
300	Lernverstärkungsmodul
320	Lernverstärkungs-Agent
330	Aktionsmodul
340	Umgebungsmodul
342	Zustandsuntermodul
343	Belohnungsuntermodul
344	Strategieuntermodul
350	Zustandsmodul
370	Belohnungsmodul
400	Ausgabemodul
450	Kalibrierungsergebnisse
900	Computerprogrammprodukt
950	Programmcode

**ZITATE ENTHALTEN IN DER BESCHREIBUNG**

*Diese Liste der vom Anmelder aufgeführten Dokumente wurde automatisiert erzeugt und ist ausschließlich zur besseren Information des Lesers aufgenommen. Die Liste ist nicht Bestandteil der deutschen Patent- bzw. Gebrauchsmusteranmeldung. Das DPMA übernimmt keinerlei Haftung für etwaige Fehler oder Auslassungen.*

**Zitierte Patentliteratur**

- CN 112632860 A [0005]
- US 20190378036 A1 [0006]
- DE 102019215530 A1 [0007]
- DE 102019208262 A1 [0008]
- EP 3825263 A1 [0009]

## Patentansprüche

1. Verfahren zum autonomen Kalibrieren eines individuellen elektrischen Antriebsstrangs (10), umfassend:

- Erstellen (S10) eines Trainingsmodells (TM) für einen elektrischen Antriebsstrang (10) von einem Lernverstärkungs-Agenten (320) mittels simulierter Beobachtungen ( $b_1, b_2, \dots, b_n$ ), wobei der Lernverstärkungs-Agent (320) einen Algorithmus des verstärkenden Lernens verwendet;
- Modifizieren (S20) des Trainingsmodells (TM) von dem Lernverstärkungs-Agenten (320) mittels realer Beobachtungen ( $b_{r1}, b_{r2}, \dots, b_{rn}$ ) eines realen idealtypischen Antriebsstrangs (10) zur Erstellung eines simulierten Modells (M) für den realen idealtypischen elektrischen Antriebsstrang (10), wobei das simulierte Modell (M) Zielzustände ( $sm_1, sm_2, \dots, sm_n$ ) enthält;
- Bestimmen (S30) zumindest eines Zustands ( $s_i$ ) eines individuellen realen elektrischen Antriebsstrangs (10) durch ein Zustandsmodul (350), wobei ein Zustand ( $s_i$ ) durch Parameter ( $p_i$ ) wie Daten und/oder Messwerte von zumindest einer Eigenschaft ( $e_i$ ) des elektrischen Antriebsstrangs (10) definiert wird,
- Übermitteln (S40) des Zustands ( $s_i$ ) an den Lernverstärkungs-Agenten (320);
- Ermitteln (S50) von Kalibrierungsergebnissen (450) für den individuellen realen elektrischen Antriebsstrang (10) von dem Lernverstärkungs-Agenten (320) durch Vergleichen des Zustands ( $s_i$ ) mit zumindest einem Zielzustand ( $sm_{ti}$ ) des simulierten Modells (M).

2. Verfahren nach Anspruch 1, wobei für das Erstellen eines Trainingsmodells (TM) für einen elektrischen Antriebsstrang (10) von einem Lernverstärkungs-Agenten (320) mittels simulierter Beobachtungen ( $b_1, b_2, \dots, b_n$ ) ein Umgebungsmodul (340) vorgesehen ist, welches zumindest ein Zustandsuntermodul (342), ein Belohnungsuntermodul (343) und ein Strategieuntermodul (344) umfasst.

3. Verfahren nach Anspruch 2, wobei von dem Zustandsuntermodul (342) Zustände ( $su_1, su_2, \dots, su_n$ ) generiert werden, die auf den simulierten Beobachtungen ( $b_1, b_2, \dots, b_n$ ) beruhen, wobei die simulierten Beobachtungen ( $b_1, b_2, \dots, b_n$ ) insbesondere die Stromstärke, die Spannung, das Drehmoment und/oder die Drehzahl eines Elektromotors oder/oder den Ladezustand einer Batterie des elektrischen Antriebsstrangs (10) umfassen.

4. Verfahren nach einem der Ansprüche 1 bis 3, wobei die realen Beobachtungen ( $b_{r1}, b_{r2}, \dots, b_{rn}$ ) gemessene Werte von Parametern ( $p_i$ ) von einer Eigenschaft ( $e_i$ ) des realen idealtypischen Antriebsstrangs (10) darstellen, die von Sensoren

ermittelt werden oder in einer Datenbank (250) gespeichert sind.

5. Verfahren nach Anspruch 1, wobei das Ermitteln von Kalibrierungsergebnissen die folgenden Verfahrensschritte umfasst:

- Auswählen einer Berechnungsfunktion ( $f_i$ ) und/oder einer Aktion ( $a_i$ ) basierend auf einer Richtlinie für einen Zustand ( $s_i$ ) für die Modifizierung zumindest eines Parameters ( $p_i$ ) von dem Lernverstärkungs-Agenten (320);
- Berechnen eines modellierten Wertes für die Eigenschaft ( $e_i$ ) mittels des modifizierten Parameters ( $p_i$ );
- Berechnen eines neuen Zustands ( $s_{i+1}$ ) von einem Umgebungsmodul (340) aufgrund des modellierten Wertes für die Eigenschaft ( $e_i$ );
- Vergleichen des neuen Zustands ( $s_{i+1}$ ) mit dem Zielzustand ( $sm_t$ ) und Zuordnen einer Abweichung ( $\Delta$ ) für das Vergleichsergebnis in dem Zustandsmodul (350);
- Ermitteln einer Belohnung ( $r_i$ ) von einem Belohnungsmodul (370) für das Vergleichsergebnis;
- Anpassen der Richtlinie des Lernverstärkungs-Agenten (320) basierend auf der Belohnung ( $r_i$ ), wobei bei einer Konvergenz der Richtlinie die optimale Aktion ( $a_j$ ) für den berechneten Zustand ( $s_j$ ) zurückgegeben wird, und bei einer Nicht-Konvergenz der Richtlinie eine weitere Berechnungsfunktion ( $f_j$ ) und/oder eine weitere Aktion ( $a_{j+1}$ ) für einen Zustand ( $s_{j+1}$ ) mit einer Modifizierung zumindest eines Parameters ( $p_j$ ) von dem Lernverstärkungs-Agenten (320) ausgewählt wird, solange bis der Zielzustand ( $sm_t$ ) erreicht ist.

6. Verfahren nach einem der Ansprüche 1 bis 5, wobei Sensoren und/oder Messvorrichtungen zur Bestimmung der Parameter ( $p_i$ ) von den Eigenschaften ( $e_i$ ) des individuellen elektrischen Antriebsstrangs (10) vorgesehen sind.

7. Verfahren nach einem der Ansprüche 1 bis 6, wobei eine positive Aktion (A+), die den Wert für einen Parameter ( $p_i$ ) erhöht, eine neutrale Aktion (A0), bei der der Wert des Parameters ( $p_i$ ) gleichbleibt, und eine negative Aktion (A-), bei der sich der Wert des Parameters ( $p_i$ ) verringert, vorgesehen sind.

8. Verfahren nach einem der vorangehenden Ansprüche 1 bis 7, wobei das Belohnungsmodul (370) eine Datenbank oder Matrix für die Bewertung der Aktionen ( $a_i$ ) umfasst.

9. Verfahren nach einem der Ansprüche 1 bis 8, wobei der zumindest eine Algorithmus des Lernverstärkungs-Agenten (320) als Markow-Entscheidungsprozess, Temporal Difference Learning (TD-Learning), Q-Learning, SARSA, Monte-Carlo-Simulation oder Actor-Critic ausgebildet ist.

10. Ein System (100) zum autonomen Kalibrieren eines individuellen elektrischen Antriebsstrangs (10), mit einem Eingabemodul (200), einem Lernverstärkungsmodul (300) und einem Ausgabemodul (400), wobei das Lernverstärkungsmodul (300) einen Lernverstärkungs-Agenten (320), der einen Algorithmus des verstärkenden Lernens verwendet, ein Aktionsmodul (330), ein Umgebungsmodul (340), ein Zustandsmodul (350) und ein Belohnungsmodul (370) umfasst; wobei der Lernverstärkungs-Agent (320) ausgebildet ist, ein Trainingsmodell (TM) für einen elektrischen Antriebsstrang (10) mittels simulierter Beobachtungen ( $b_1, b_2, \dots, b_n$ ) zu erstellen, und das Trainingsmodell (TM) mittels realer Beobachtungen ( $b_{r1}, b_{r2}, \dots, b_{rm}$ ) eines realen idealtypischen Antriebsstrangs (10) zu modifizieren zur Erstellung eines simulierten Modells (M) für den realen idealtypischen elektrischen Antriebsstrang (10), wobei das simulierte Modell (M) Zielzustände ( $sm_1, sm_2, \dots, sm_n$ ) enthält; wobei das Zustandsmodul (350) ausgebildet ist, zumindest einen Zustand ( $s_i$ ) eines individuellen realen elektrischen Antriebsstrangs (10) zu bestimmen, wobei ein Zustand ( $s_i$ ) durch Parameter ( $p_i$ ) wie Daten und/oder Messwerte von zumindest einer Eigenschaft ( $e_i$ ) des elektrischen Antriebsstrangs (10) definiert wird, und den Zustand ( $s_i$ ) an den Lernverstärkungs-Agenten (320) zu übermitteln; und wobei der Lernverstärkungs-Agent (320) ausgebildet ist, Kalibrierungsergebnisse (450) für den individuellen realen elektrischen Antriebsstrang (10) durch Vergleichen des Zustands ( $s_i$ ) mit zumindest einem Zielzustand ( $sm_{ti}$ ) des simulierten Modells (M) zu ermitteln.

11. System (100) nach Anspruch 10, wobei das Umgebungsmodul (340) zumindest ein Zustandsuntermodul (342), ein Belohnungsuntermodul (343) und ein Strategieuntermodul (344) umfasst.

12. System (100) nach Anspruch 11, wobei das Zustandsuntermodul (342) ausgebildet ist, Zustände ( $su_1, su_2, \dots, su_n$ ) zu generieren, die auf den simulierten Beobachtungen ( $b_1, b_2, \dots, b_n$ ) beruhen, wobei die simulierten Beobachtungen ( $b_1, b_2, \dots, b_n$ ) insbesondere die Stromstärke, die Spannung, das Drehmoment und/oder die Drehzahl eines Elektromotors und/oder den Ladezustand einer Batterie des elektrischen Antriebsstrangs (10) umfassen.

13. System (100) nach einem der Ansprüche 10 bis 12, wobei die realen Beobachtungen ( $b_{r1}, b_{r2}, \dots, b_{rm}$ ) gemessene Werte von Parametern ( $p_i$ ) von einer Eigenschaft ( $e_i$ ) darstellen, die von Sensoren ermittelt werden oder die in einer Datenbank (250) gespeichert sind.

14. System (100) nach einem der Ansprüche 10 bis 13, wobei Sensoren und/oder Messvorrichtungen zur Bestimmung der Parameter ( $p_i$ ) von den

Eigenschaften ( $e_i$ ) des individuellen elektrischen Antriebsstrangs (10) vorgesehen sind.

15. Computerprogrammprodukt (900), umfassend einen ausführbaren Programmcode (950), der so konfiguriert ist, dass er bei seiner Ausführung das Verfahren nach einem der Ansprüche 1 bis 9 ausführt.

Es folgen 2 Seiten Zeichnungen

Anhängende Zeichnungen

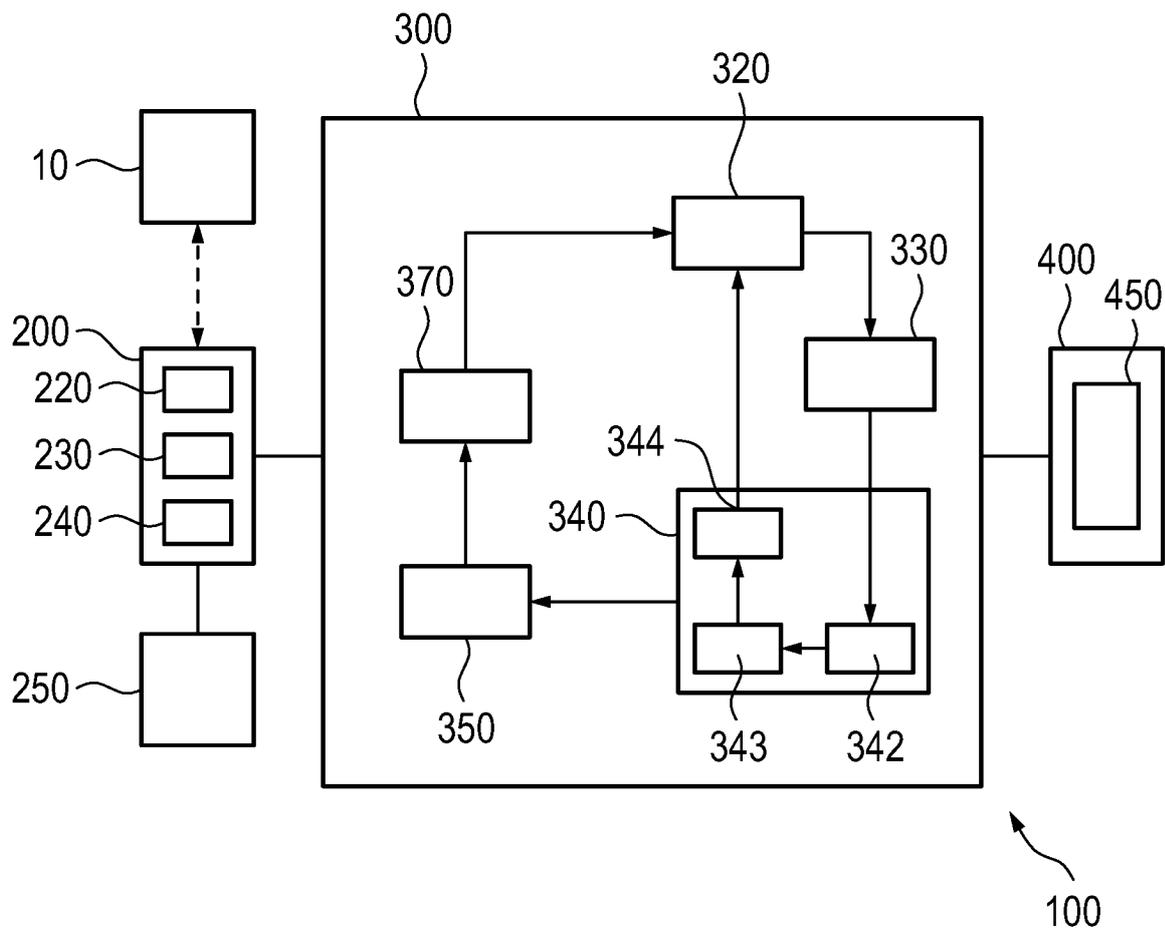
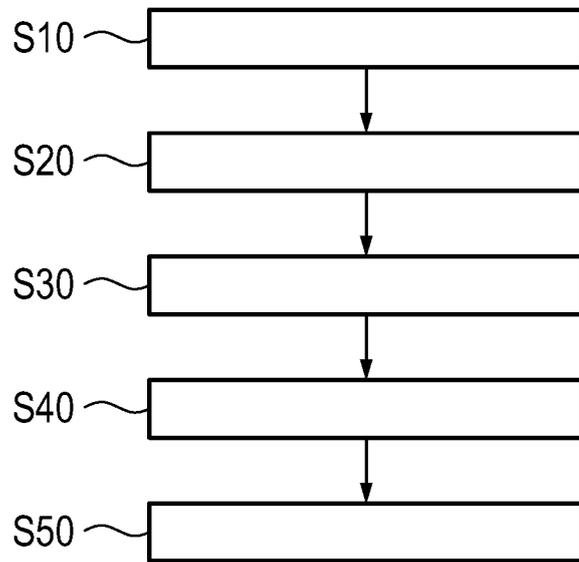
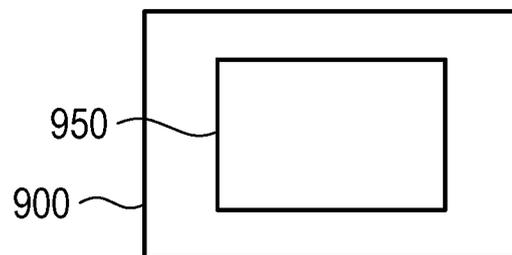


Fig. 1



**Fig. 2**



**Fig. 3**