

Petitioner's Presentation

**Digital Global Systems, Inc. v.
DeepSig, Inc.**

IPR2024-01358 (U.S. Pat. 11,018,704)

December 16, 2025

The '704 Patent

US011018704B1

(12) **United States Patent**
O'Shea et al.

(10) **Patent No.:** US 11,018,704 B1
(45) **Date of Patent:** May 25, 2021

(54) **MACHINE LEARNING-BASED NONLINEAR PRE-DISTORTION SYSTEM**
USPC 455/114.3, 63.1, 501, 67.13
See application file for complete search history.

(71) Applicant: **DeepSig Inc.**, Arlington, VA (US)
(72) Inventors: **Timothy James O'Shea**, Arlington, VA (US); **James Shea**, Arlington, VA (US)
(73) Assignee: **DeepSig Inc.**, Arlington, VA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **16/806,247**
(22) Filed: **Mar. 2, 2020**

Related U.S. Application Data
(63) Continuation of application No. 15/955,485, filed on Apr. 17, 2018, now Pat. No. 10,581,469.
(60) Provisional application No. 62/486,200, filed on Apr. 17, 2017.

(51) **Int. Cl.**
H04B 1/04 (2006.01)
G06N 3/08 (2006.01)
G06N 3/04 (2006.01)
H04W 88/08 (2009.01)

(52) **U.S. Cl.**
CPC *H04B 1/0475* (2013.01); *G06N 3/04* (2013.01); *G06N 3/08* (2013.01); *H04W 88/08* (2013.01)

(58) **Field of Classification Search**
CPC *H04B 1/0475*; *G06N 3/04*; *G06N 3/08*; *H04W 88/08*

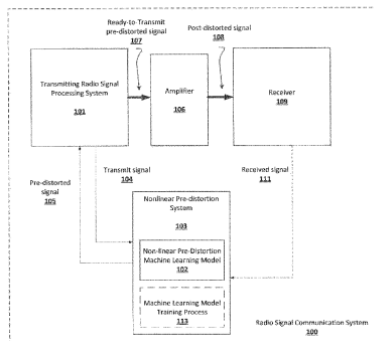
References Cited
U.S. PATENT DOCUMENTS
6,141,390 A 10/2000 Ceva
7,289,773 B2 10/2007 Heathwaite
7,372,918 B2 5/2008 Muller
9,245,072 B2* 1/2016 Jia G06F 30/367
9,307,619 B2* 7/2016 Loshkin H03F 1/3258
9,444,412 B2* 9/2016 Loshkin H03G 3/3042
10,075,324 B2 9/2018 Zhao
10,581,469 B1* 3/2020 O'Shea H04B 1/0475
2017/0257251 A1* 9/2017 Zhao H03F 1/3258

* cited by examiner

Primary Examiner — Ankur Jain
(74) **Attorney, Agent, or Firm** — Fish & Richardson P.C.

ABSTRACT
(57) Methods, systems, and apparatus, including computer programs encoded on computer storage media, for correcting distortion of radio signals. A transmit radio signal corresponding to an output of a transmitting radio signal processing system is obtained. A pre-distorted radio signal is then generated by processing the transmit radio signal using a nonlinear pre-distortion machine learning model. The nonlinear pre-distortion machine learning model includes model parameters and at least one nonlinear function to correct radio signal distortion or interference. A transmit output radio signal is obtained by processing the pre-distorted radio signal through the transmitting radio signal processing system. The transmit output radio signal is then transmitted to one or more radio receivers.

24 Claims, 12 Drawing Sheets



- Earliest priority date of April 17, 2017
- Predistortion and neural networks were known before the critical date. Petition, 4-9.
- The NPDMLM being “associated with values corresponding to one or more deployment scenarios” as well as “cellular communications” and a “base station” was not novel either. See Ex-1002 ('704 patent), 44, 63-64.

Digital Global Systems's Grounds

Ground	Basis	Claims	References
1	§103	1-3, 7-8, 9-11, 15-16, 17-19, 23-24	Jüschke + POSITA
2	§103	1-24	Ground 1 + Holt
3	§103	2-6, 10-14, 18-22	Ground 2 + Dzierwa

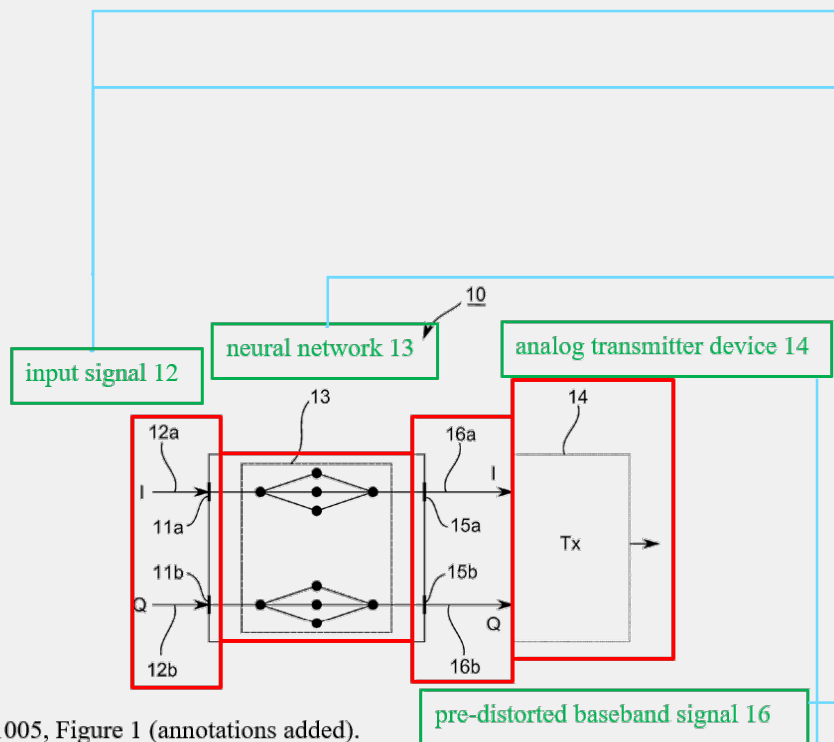
(Petition, 3.) Grounds 1-3 are supported by the testimony of Dr. Ding. Ex-1003.

DeepSig's POR is:

- Limited to [1B.1] & [1B.2] and similar limitations
- Limited to claim 4 and similar claims
- Unsupported by expert testimony

Jüschke renders claim 1 obvious

(and similar claims 9 and 17)



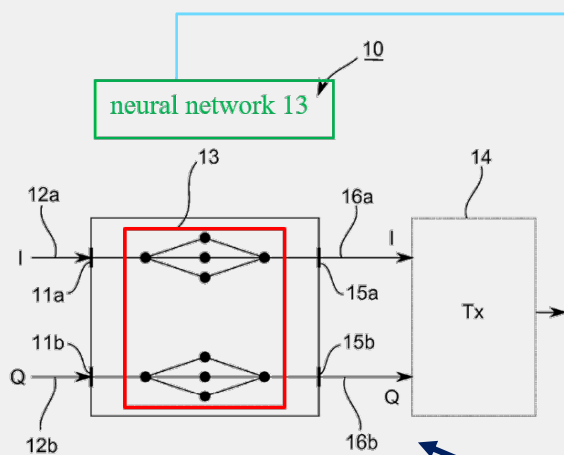
Ex-1005, Figure 1 (annotations added).

Claim 1	Limitation
[1P]	A method for correcting distortion of radio signals in a cellular communications system:
[1A]	obtaining, at a base station in the cellular communications system, a <u>transmit radio signal</u> that is prepared for transmitting to one or more radio receivers associated with the cellular communications system;
[1B.1]	accessing a <u>nonlinear pre-distortion machine learning model</u> associated with the base station,
[1B.2]	wherein the <u>nonlinear pre-distortion machine learning model</u> includes model parameters associated with values corresponding to one or more deployment scenarios in the cellular communications system,
[1B.3]	and at least one nonlinear function to correct radio signal distortion;
[1C]	generating a <u>pre-distorted radio signal</u> by processing the transmit radio signal using the nonlinear pre-distortion machine learning model;
[1D]	obtaining a <u>transmit output radio signal</u> by processing the pre-distorted radio signal through a transmitting radio signal processing stage of the base station; and
[1E]	<u>transmitting the transmit output radio signal</u> to the one or more radio receivers.

Petition, 20-37 (claim 1).

Jüschke renders [1B.1] obvious

Jüschke explicitly teaches that its radio transmitter (and NN) can be used for base stations. See also Ding Declaration, ¶¶122, 131.



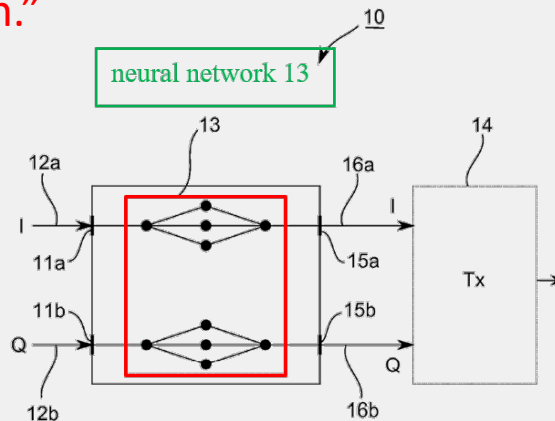
Claim 1	Limitation
[1B.1]	accessing a <u>nonlinear pre-distortion machine learning model associated with the base station,</u>

Ex-1005, Figure 1 (annotations added).

[0006] Therefore, it is desirable to have a more flexible radio transmitter that can support as many standards and signals as possible. Beside the interest on increased flexibility from the outside view, a high flexibility within the transceiver system has the advantage, that different parts of the system can be exchanged and adapted easily, like the flexibility to use the digital part with different analog components without reconfiguration. This is not only true for macro-cell base stations but also for small-cell base stations.

[1B.1] PO's Argument Implemented ≠ Associated is Incorrect and Unsupported

- Patent Owner incorrectly argues that “Petitioner plainly fails to identify any association between a base station in which a transmitter is located and Jüschke’s neural network.” POR, 32.
- Patent Owner provides no extrinsic evidence or expert testimony to support this.
- Jüschke clearly teaches that NN 13 within its “flexible radio transmitter” can be associated with a “base station.”



Ex-1005, Figure 1 (annotations added).

[0006] Therefore, it is desirable to have a more flexible radio transmitter that can support as many standards and signals as possible. Beside the interest on increased flexibility from the outside view, a high flexibility within the transceiver system has the advantage, that different parts of the system can be exchanged and adapted easily, like the flexibility to use the digital part with different analog components without reconfiguration. This is not only true for macro-cell base stations but also for small-cell base stations.

[1B.1] PO's Argument Implemented ≠ Associated is Incorrect and Unsupported

- Patent Owner's implemented ≠ associated with argument is pure attorney argument, unsupported by extrinsic evidence, expert testimony, or a claim construction.
- The Board correctly rejected this argument in granting institution.

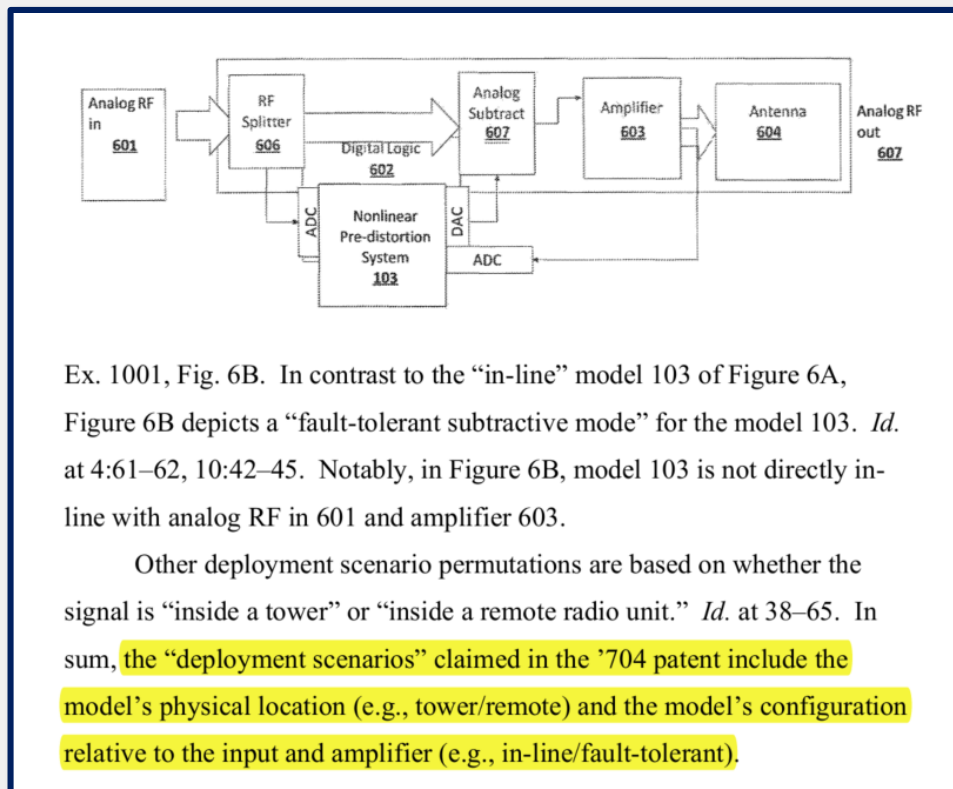
1. “nonlinear pre-distortion machine learning model associated with the base station”

Petitioner asserts that Jüschke discloses a system “for mobile communications” that uses “base stations.” Pet. 22–23 (citing Ex. 1005 ¶¶ 2, 6). Patent Owner disputes whether Jüschke discloses the model associated with the base station as claimed, focusing on the words “associated with.” Prelim. Resp. 26–28. Patent Owner argues that Petitioner’s position is that the model is “implemented within” the base station, but that the claim requires the model to be “associated with” the base station. *Id.* at 27. On this record, with no claim construction offered, we fail to see how Patent Owner construes a model “implemented within,” i.e., used in, a base station to not also be “associated with” the base station. That is, it would seem that a model used in a base station is associated with that base station. Accordingly, we do not find Patent Owner’s argument persuasive on this preliminary record.

Institution Decision, 11-12.

Jüschke renders [1B.2] obvious

- As the Board correctly found, the “deployment scenarios” include the NN model’s physical location and configuration relative to the input and amplifier.
- I.e., these are hardware configurations with respect to the NN model.

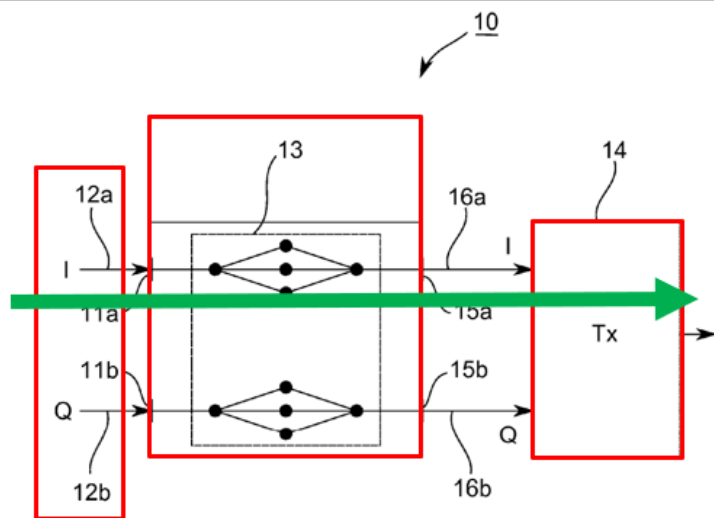


Institution Decision, 8.

Jüschke renders [1B.2] obvious

Jüschke explicitly teaches that its model parameters are correspond to deployment scenarios exactly matching those in the '704 patent.

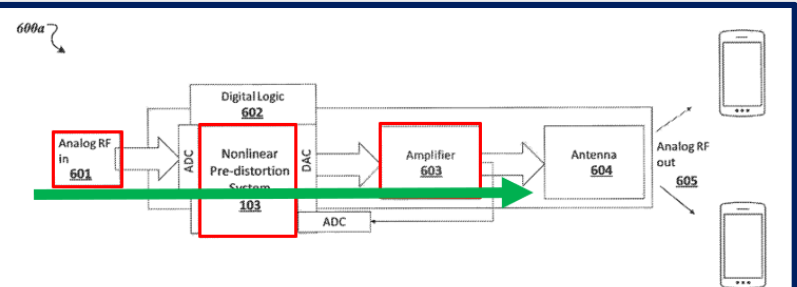
See *also* Ding Declaration, ¶¶141-48 ([1B.2]), 185-89 ([3A]/[3C]).



Jüschke's flexible transmitter hardware with pre-distortion NN 13 inline between input transmit signal 12 and power amplifier in transmitter device 14.

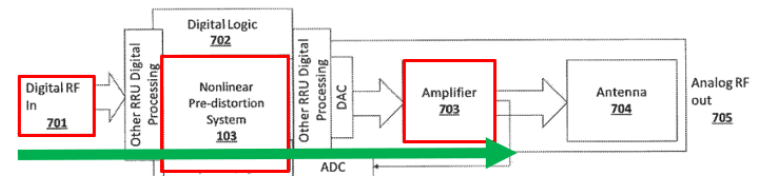
Jüschke, Figure 1 (annotations added).

Claim 1	Limitation
[1B.2]	wherein the nonlinear pre-distortion machine learning model includes model parameters associated with values corresponding to one or more <u>deployment scenarios in the cellular communications system</u> ,



Tower-mounted amplifier with non-linear pre-distortion system 103 inline between input 601 and amplifier 603.

FIG. 6A



Remote radio unit with non-linear pre-distortion system 103 inline between input 701 and amplifier 703.

FIG. 7A

'704 patent, Figures 6A, 7A (annotations added).

Jüschke renders [1B.2] obvious

- The Board correctly found a “**striking**” resemblance between Jüschke Figure 1 and ’704 patent Figure 6A.

2. “model parameters associated with values corresponding to one or more deployment scenarios”

Petitioner asserts that the “model parameters associated with values” are the “weight and/or bias parameters” of the model. Pet. 26–27. Petitioner further asserts that the configuration of the model in Jüschke is the same configuration as the “in-line” deployment scenario described in the ’704 patent. *Id.* at 26–29. Patent Owner argues that “Jüschke does not disclose any association of the weight or bias parameters, which are used to approximate the transmitter device’s inverse transfer function, to any deployment scenario.” Prelim. Resp. 15; *see also id.* at 12–16 (entire argument). However, the in-line deployment scenario depicted in Figure 6a of the ’704 patent bears a striking resemblance to the arrangement depicted in Figure 1 of Jüschke. In particular, both figures show the analog radio signal fed directly into the model and then directly into the amplifier. Compare Ex. 1001, Fig. 6A with Ex. 1005, Fig. 1. As we explained in our preliminary claim construction section, the claimed “deployment scenarios” include the model’s configuration relative to the input and amplifier, such as “in-line.” Thus, on the present record, we do not find Patent Owner’s arguments regarding the “deployment scenarios” limitation persuasive.

Institution Decision, 12-13.

[1B.2] PO Incorrectly Adds “Use Dependent” to the Claim Language

- Patent Owner incorrectly argues that [1B.2] requires the “deployment scenarios” must be “use dependent.” POR, 26.
- There is no such “use dependent” requirement; only that the parameters correspond to “one or more deployment scenarios.”

Claim 1	Limitation
[1B.2]	wherein the nonlinear pre-distortion machine learning model includes model parameters associated with values corresponding to one or more deployment scenarios in the cellular communications system,

[1B.2] PO Incorrectly Adds “Use Dependent” to the Claim Language

- The Board correctly rejected this “use dependent” v. inverse transfer function argument.
- Jüschke’s model ***functioning in given deployment scenario*** teaches that model’s ***parameters and values corresponding to said scenario***.

2. “model parameters associated with values corresponding to one or more deployment scenarios”

Petitioner asserts that the “model parameters associated with values” are the “weight and/or bias parameters” of the model. Pet. 26–27. Petitioner further asserts that the configuration of the model in Jüschke is the same configuration as the “in-line” deployment scenario described in the ’704 patent. *Id.* at 26–29. Patent Owner argues that “Jüschke does not disclose any association of the weight or bias parameters, which are used to approximate the transmitter device’s inverse transfer function, to any deployment scenario.” Prelim. Resp. 15; *see also id.* at 12–16 (entire argument). However, the in-line deployment scenario depicted in Figure 6a of the ’704 patent bears a striking resemblance to the arrangement depicted in Figure 1 of Jüschke. In particular, both figures show the analog radio signal fed directly into the model and then directly into the amplifier. Compare Ex. 1001, Fig. 6A with Ex. 1005, Fig. 1. As we explained in our preliminary claim construction section, the claimed “deployment scenarios” include the model’s configuration relative to the input and amplifier, such as “in-line.” Thus, on the present record, we do not find Patent Owner’s arguments regarding the “deployment scenarios” limitation persuasive.

Patent Owner similarly argues that the parameters and values of Jüschke do not meet the claim language. Prelim. Resp. 18–21. For example, Patent Owner argues that Jüschke’s parameters are “generic” and that Petitioner has not shown “that the parameters themselves are selected for the model to correspond to one or more deployment scenarios.” *Id.* at 21. However, as we explained above, Petitioner has shown preliminarily that Jüschke teaches or suggests a model that is deployed in-line in the base station, i.e., a “deployment scenario” as claimed. Given that the model is functioning in a claimed “deployment scenario,” then it stands to reason that its parameters and values correspond to that scenario. To the extent Patent Owner is asserting that the model must have many configurations, which can be selected once placed in a given deployment scenario, we note that the claims are not so limited. Ex. 1001, 14:25. The plain meaning of “one or more” includes “one,” i.e., one and no more. Thus, on the present record, we do not find Patent Owner’s arguments regarding the “parameters” limitation persuasive.

[1B.2] PO Incorrectly Adds a Selection Step to the Claim Language

- Patent Owner incorrectly argues that [1B.2] requires an extra “selection” step requiring the parameters be selected from a group of two or more sets of parameters corresponding to different “deployment scenarios.” POR, 25-26.
- Patent Owner further incorrectly implies that “one or more” means more than one. POR, 25.
- There is no such “selection” or “more than one” requirement; only that the parameters correspond to “one or more deployment scenarios.”

Claim 1	Limitation
[1B.2]	wherein the nonlinear pre-distortion machine learning model includes model parameters associated with values corresponding to one or more deployment scenarios in the cellular communications system,

[1B.2] PO Incorrectly Adds a Selection Step to the Claim Language

- The Board correctly rejected both PO's extra "selection step" requirement and reading of "one *or* more" as more than one.

Patent Owner similarly argues that the parameters and values of Jüschke do not meet the claim language. Prelim. Resp. 18–21. For example, Patent Owner argues that Jüschke's parameters are "generic" and that Petitioner has not shown "that the parameters themselves are selected for the model to correspond to one or more deployment scenarios." *Id.* at 21. However, as we explained above, Petitioner has shown preliminarily that Jüschke teaches or suggests a model that is deployed in-line in the base station, i.e., a "deployment scenario" as claimed. Given that the model is functioning in a claimed "deployment scenario," then it stands to reason that its parameters and values correspond to that scenario. To the extent Patent Owner is asserting that the model must have many configurations, which can be selected once placed in a given deployment scenario, we note that the claims are not so limited. Ex. 1001, 14:25. The plain meaning of "one *or* more" includes "one," i.e., one and no more. Thus, on the present record, we do not find Patent Owner's arguments

Jüschke and Holt renders claim 4 obvious

- The “*received radio signal*” only needs to “***correspond[] to a representation*** of the transmit output radio signal.”
- This can be a modeled signal.
- The “*received radio signal*” in [4B] refers back to the “*received radio signal that corresponds to a representation*” in [4A] (e.g., modeled signal).

Claim 4	Limitation
[4A]	The method of claim 2, further comprising: obtaining a received radio signal that <u>corresponds</u> to a <u>representation</u> of the transmit output radio signal received at the one or more radio receivers;
[4B]	determining a distance metric between the transmit radio signal and the received radio signal; and
[4C]	updating the values of the model parameters for the particular deployment scenario based at least on the distance metric.

(and similar claims 12 and 20)

Petitioner’s Reply, 17-18.

Jüschke and Holt renders claim 4 obvious

- Related U.S. Patent No. 11,777,540 (child to the '540 patent) further shows that the “*received radio signal*” can be a “**modeled**” signal.
- The same modeled “*received radio signal*” is used to calculated a distance metric.

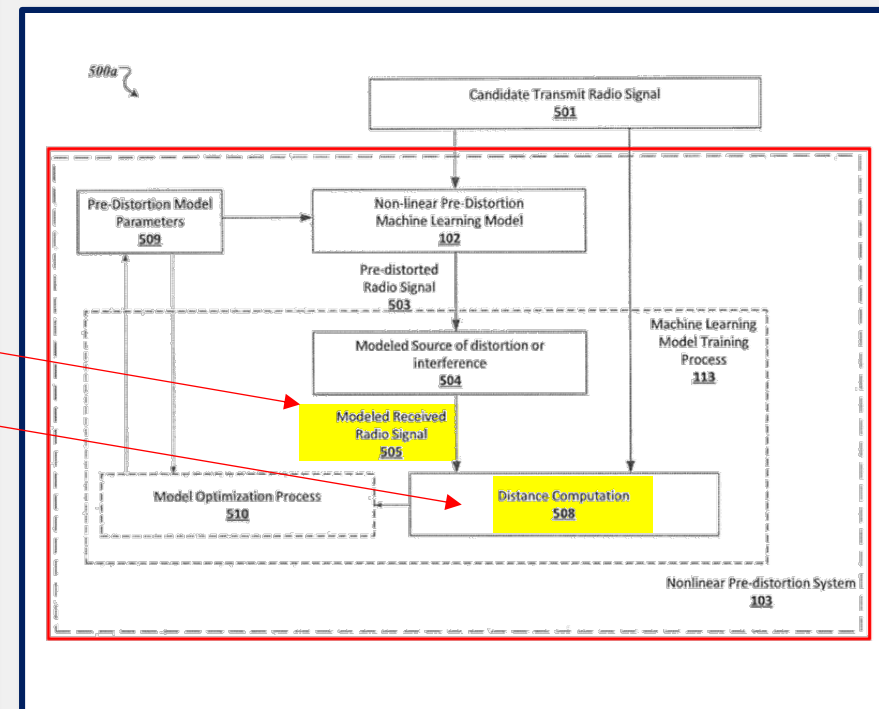
Claim 12	Limitation
'540 [12E]	updating the model parameters based on a distance metric computed between the transmit radio signal and the received radio signal.

Claim 13	Limitation
'540 [13]	The method of claim 12, wherein processing the pre-distorted radio signal comprises generating the received radio signal using a modeled source of distortion or interference.

Jüschke renders claim 4 obvious

- The '704 patent directly supports the use of modeled signals for calculating a distance metric between a candidate (input) signal 501 and a (modeled) received signal 505.

FIG. 5A illustrates an example configuration of a nonlinear pre-distortion system for model training. In some implementations, the nonlinear pre-distortion system 500a is the same as the nonlinear pre-distortion system 103. The nonlinear pre-distortion machine learning model in configuration 500a provides a pre-distorted radio signal 503 to a modeled source of distortion or interference 504. The distortion or interference source produces a modeled received radio signal 505 that can be used by a distance computation process 508 to compute a distance metric between the candidate transmit radio signal 501 and the modeled received radio signal 505. The system then sends the computed distance metric to a model optimization process 510. This process determines updated parameters for the machine learning model that minimize the distance metric. The goal



Ex-1001 ('704 patent), 9:37-51.

Ex-1001 ('704 patent), Figure 5A (annotations added).

Jüschke and Holt renders claim 4 obvious

- Likewise, Holt teaches calculating a first loss function (216) between an input signal (204) and a modeled output signal (214).

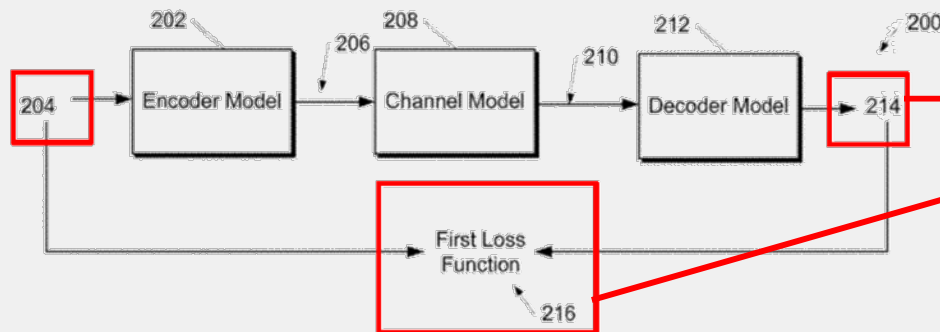


Figure 3

45 More particularly, FIG. 3 depicts a block diagram of an example loss function (e.g., first loss function 216) defined between the first set of inputs 204 and the third set of outputs 214. The first loss function 216 can describe a difference between the first set of inputs 204 and the third set of outputs 214. For example, the first loss function 216 can be a squared difference between the first set of inputs 204 and the third set of outputs 214 (e.g., a sum of a squared difference for each dimension in the first set of dimensions). However, loss functions other than a squared difference can be used as
50 well, including, for example, loss functions that include various weightings (e.g., a particular weight applied to each dimension in the first set of dimensions).

Holt, Figure 3 (annotations added).

Holt, 16:45-57.

Jüschke and Holt renders claim 4 obvious

- Likewise, Holt teaches calculating a first loss function (216) between an input signal (204) and a modeled output signal (214).

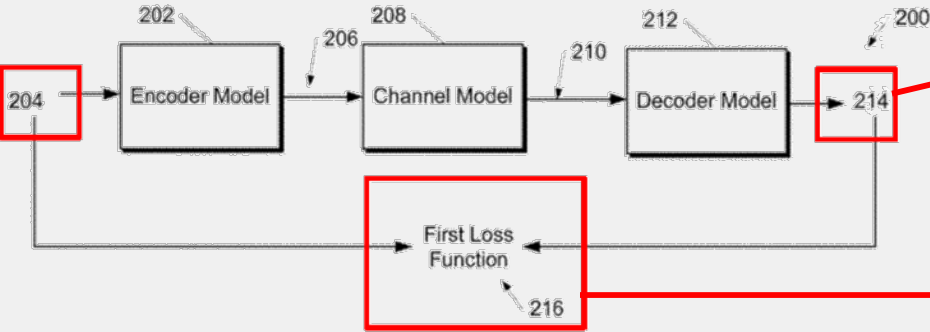


Figure 3

Holt, Figure 3 (annotations added).

Claim 4	Limitation
[4A]	The method of claim 2, further comprising: obtaining a received radio signal that corresponds to a representation of the transmit output radio signal received at the one or more radio receivers;
[4B]	determining a distance metric between the transmit radio signal and the received radio signal ; and
[4C]	updating the values of the model parameters for the particular deployment scenario based at least on the distance metric.

Jüschke and Holt renders claim 4 obvious

- Likewise, Holt teaches calculating a first loss function (216) between an input signal (204) and a modeled output signal (214).

The first loss function can be backpropagated through the decoder model while modifying the decoder model to train the decoder model, for example, by modifying at least one weight associated with the decoder model. After backpropagating the first loss function through the decoder model, the first loss function can continue to be backpropagated through the channel model without modifying the channel model. After backpropagating the first loss function through the channel model, the first loss function can be backpropagated through the encoder model while modifying the encoder model to train the encoder model, for example, by modifying at least one weight associated with the encoder model. As such, both the encoder and decoder models can be trained relative to the inputs and outputs of the channel model.

Claim 4	Limitation
[4A]	The method of claim 2, further comprising: obtaining a received radio signal that corresponds to a representation of the transmit output radio signal received at the one or more radio receivers;
[4B]	determining a distance metric between the transmit radio signal and the received radio signal ; and
[4C]	updating the values of the <u>model parameters</u> for the particular deployment scenario <u>based at least on the distance metric</u> .

Ex-1006 (Holt), 5:67-6:14.

Jüschke and Holt renders claim 4 obvious

- Jüschke and Holt teaching a training an encoder model using a modeled signal and rendering claim 4 obvious is consistent with the Board's previous rulings on the related patents.

the nonlinear pre-distortion system"). The portions of Holt cited by Petitioner teach training an encoder (i.e., transmit) model based on a loss function computed using a signal obtained from a model, not from a

IPR2025-00049 (U.S. 10,581,469), Paper 7, 12-13.

4. Claim 13

Claim 13 depends from claim 12, and recites "wherein processing the pre-distorted signal comprises generating the received radio signal using a modeled source of distortion or interference." Ex. 1001, 15:62-65. Petitioner argues that "Holt Figure 3 shows channel model 208 implemented with encoder/transmitter 202, decoder/receiver 212, and loss function 216." Pet. 73 (citing Ex. 1006, Fig. 3). According to Petitioner, a person of ordinary skill in the art "would have understood that the combined Jüschke/Holt model training system could implement Holt's channel model trained using a modeled form of noise." *Id.* at 74 (citing Ex. 1003 ¶ 273).

In other words, for claim 13, Petitioner relies on Holt's embodiments that include a channel model, rather than the embodiments that include an actual communication channel. Pet. 71-74. As discussed above, Holt teaches using a loss function to "train[] the encoder model and decoder model when the channel model is structurally provided between the encoder model and decoder model." Ex. 1006, 5:52-56 (emphasis added); *see id.* at code (57), 9:61-63, 13:4-20, 16:60-17:10; Section II.D.1.b. Thus, Petitioner's arguments for claim 13 do not suffer from the same deficiencies discussed above for claims 4 and 12. *See* Sections II.D.1, II.D.3.

IPR2025-00194 (U.S. 11,777,540), Paper 10, 15.

Claim 4	Limitation
[4A]	The method of claim 2, further comprising: obtaining a received radio signal that <u>corresponds to a representation of the transmit output radio signal received at the one or more radio receivers;</u>
[4B]	determining a distance metric between the transmit radio signal and the received radio signal; and
[4C]	updating the values of the <u>model parameters for the particular deployment scenario based at least on the distance metric.</u>

See Petitioner's Reply, 23 n.6 (pointing to '540 claim 13 and IPR2025-00194).